UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

Genetic characterisation of an African elephant (*Loxodonta africana*) population: the role of genetic relatedness in male social groups

Teresa Leonor Nunes dos Santos Nobre dos Santos

Mestrado em Biologia Evolutiva e do Desenvolvimento

Dissertação orientada por:
Doutora Hannah Mumby
Professor Doutor Carlos Fernandes

2017

Acknowledgments

This thesis could not have been finished without the help and support of countless people, and I am thankful to each and every one of them.

To Dr. Hannah Mumby, for showing me the beauty of elephants, and for the trust and constant encouragement to go further.

To Professor Carlos Fernandes, whose door was always open when I could not make sense of the data.

To the rest of the Bull Network Project team, for being an absolute delight to work with, and especially to Dr. Derek Murphy who calculated the social networks and social association vs. genetic relatedness correlations presented herein.

To the Elephants Alive team, and especially Dr. Michelle Henley, for collecting the samples and field data.

To Dr. Sam Wasser's team, for providing us with Kruger elephant samples.

To the Molecular Ecology Lab at the University of Sheffield, for all their invaluable help in the lab, including but not limited to Prof. Terry Burke, Dr. Natalie dos Remédios, Dr. Gavin Hosborough, Rachel Tucker, Dr. Helen Hipperson, and Dr. Deborah Dawson.

To my friends, old and new, for encouraging me to step towards the unknown, for walking by my side, and for not staring at me oddly when I spoke of faecal samples. A special thank you to my "thesis colleagues", Rita Bastos and Rodrigo Costa; to my UK buddies, Lucy Garrett, Natalie dos Remédios, Jamie Hutchinson, Jamie Thompson, Michael Hutchinson, Amy Leedale, Carly Lynsdale, and Cristina Pecheanu; and to my favourite companion, Diana Sousa.

To my parents, for having always believed and fought for me, even when I did not, and to my little brother, who never forgets to do my head in.

And to the elephants, for being so utterly fascinating.

Resumo

O elefante da savana africano (*Loxodonta africana*) é uma espécie altamente social, na qual é observável a separação de machos e fêmeas quando os indivíduos atingem a maturidade sexual. O número de elefantes desta espécie está a diminuir por todo o continente africano, devido à ocorrência histórica e actual de abate de indivíduos para controlo populacional em zonas de maior densidade de elefantes e de caça ilegal, bem como graças à crescente fragmentação do habitat e ocupação humana. Isto potencia o decréscimo da variabilidade genética, que, por sua vez, pode ter impactos negativos na sobrevivência das populações. De facto, aumenta a probabilidade de ocorrer um *bottleneck* genético, depressão de consanguinidade, e uma diminuição do tamanho populacional efectivo. Este problema pode ser exacerbado quando se tem em conta a presença de barreiras físicas e de populações humanas que impedem a migração de animais e, consequentemente, o fluxo genético. O isolamento populacional aumenta os efeitos da deriva genética e promove a diferenciação genética entre populações.

Adicionalmente, os elefantes mais velhos são o principal alvo de caçadores furtivos por apresentarem maiores presas. Estes indivíduos não só têm posições mais centrais em redes sociais, como actuam como um repositório de conhecimento, tanto social como ecológico, para os elefantes mais novos. Remover estes indivíduos pode levar a alterações comportamentais duradouras, que podem mesmo conduzir à redução do *fitness* reprodutivo.

Devido à importância da sociabilidade nesta espécie e do impacto que poderá ter sobre a sua sobrevivência, urge compreender como a semelhança genética e o comportamento estarão associados nos elefantes. É necessário, em particular, estudar este parâmetro entre machos. Estes não só são caçados com maior frequência devido ao tamanho das suas presas, como apresentam comportamentos conflituosos para com humanos. Além disto, enquanto os grupos de fêmeas estão altamente estudados a nível social e genético, ainda pouco se conhece sobre os laços sociais entre machos. Os grupos de fêmeas são matriarcais, formados por até vinte adultos aparentados e as suas crias. Os machos, ao atingirem a adolescência, dispersam do grupo natal.

Os elefantes macho são geralmente solitários, pelo que a constituição de tais grupos é altamente variável em termos do número de indivíduos, da sua coesão, duração e composição. No entanto, não existem estudos que explicitem com alto grau de certeza o que determina a sua formação e manutenção, nem existem descrições completas das associações entre machos. Também não é sabido se estas associações têm impacto sobre a sobreposição de áreas vitais. De facto, apesar de os machos não serem territoriais, tendem a evitar outros machos durante a época de cio. É possível que demonstrem diferentes níveis de tolerância a outros machos dependendo da relação social entre eles estabelecida. Caso o parentesco genético seja um factor pertinente para a formação destas relações, espera-se que traga tais benefícios aos mesmos, como a diminuição da agressividade entre machos e um aumento de sucesso reprodutivo.

Neste estudo, temos como objectivo primário estudar a genética populacional de uma população de elefantes da savana na África do Sul, mais especificamente nas Associated Private Nature Reserves (APNR), um conjunto de áreas privadas adjacentes ao Parque Nacional Kruger (PNK), na África do Sul, e ao Parque Nacional do Limpopo, em Moçambique. Até à década de 1990, as APNR encontravam-se cercadas, não havendo migração de indivíduos para dentro ou fora da região. Como tal, e tendo em conta a longa longevidade dos elefantes, é possível que ainda sejam visíveis os efeitos deste isolamento a nível genético. Por outro lado, observações efectuadas no campo indicam que se tem dado um aumento populacional. Porém, ainda não foi feita uma análise genética que permita compreender se o aumento populacional está a ser acompanhado por um incremento de diversidade genética. Caso este seja o caso e não seja detectada elevada consanguinidade ou um *bottleneck* genético, a população da APNR poderá ser ideal para futuros estudos relativos ao impacto do parentesco genético na formação e manutenção de associações sociais entre machos, visto não ser esperado um grande impacto da caça furtiva na

população. Logo, é expectável que as associações e comportamentos demonstrados sejam representativos de grupos sociais não perturbados.

Para realizar este estudo, foram genotipadas amostras fecais não invasivas de 80 fêmeas e 168 machos da APNR, dos quais 68 são juvenis e 180 são adultos, com um painel de 18 marcadores de microssatélites autossómicos, suplementados por três marcadores de microssatélites para confirmação da identificação sexual. Comparámos ainda os resultados obtidos com 46 elefantes do PNK. Analisámos o nível de diversidade genética e estimámos o tamanho populacional efectivo (Ne = 394.4). Não encontrámos evidências para a existência de um *bottleneck* genético ou de consanguinidade. Também não foi observável estrutura genética, quer entre os elefantes das APNR e PNK, quer dentro de cada população, para ambos os sexos. O nível de diferenciação genética entre a APNR e o PNK foi igualmente baixo (Fst = 0.0038).

Calculámos ainda o nível de parentesco entre os machos e criámos uma rede genética. De seguida, mapeámos esta rede sobre índices preliminares de associação entre 16 machos, não tendo sido obtida uma correlação entre associação social e parentesco genético. No entanto, obtivemos uma correlação entre a proximidade de amostras e parentesco genético, sendo que, a distâncias inferiores a 500 metros, os animais eram mais aparentados geneticamente do que o esperado. Tal demonstra que indivíduos aparentados apresentam pelo menos sobreposição parcial de áreas vitais. Note-se que a componente social deste trabalho carece de uma maior quantidade de observações e de um maior número de machos observados de modo a permitir tirar ilações robustas dos resultados obtidos. Esperamos, assim, obter mais dados sociais no futuro.

Os resultados obtidos indicam que esta população apresenta uma boa saúde ao nível da diversidade genética. No entanto, estes resultados poderão advir em parte da existência de migração entre regiões, que poderá introduzir novos alelos na população, e consequentemente encobrir a assinatura genética do isolamento populacional histórico, bem como aumentar o nível de variabilidade genética e o tamanho populacional efectivo. Igualmente, a migração poderá ter sido suficiente para diminuir o nível de diferenciação genética entre os elefantes amostrados nas APNR e no PNK. Como tal, é importante que a população continue a ser protegida, de modo a que a caça furtiva no sul de África não leve à redução da variabilidade genética e à disrupção de grupos sociais. A conservação desta população torna-se ainda mais importante quando se tem em conta o acentuado decréscimo populacional da espécie ao nível do continente africano.

Este estudo piloto caracteriza geneticamente uma população de elefantes sul-africana que demonstra a importância do fluxo genético para manter a diversidade genética e mitigar os efeitos nocivos da caça furtiva e do abate de animais. Como tal, a manutenção de conectividade entre parques nacionais e reservas naturais é vital para garantir a migração de elefantes entre populações. Este estudo é ainda um primeiro passo para uma caracterização pormenorizada do modo como o parentesco genético pode explicar tais comportamentos de machos como a escolha de indivíduos com quem se associar e padrões de dispersão.

Palavras-chave: *Loxodonta africana*, genética populacional, estrutura genética, comportamento social

Abstract

The African savannah elephant (*Loxodonta africana*) is a highly social species that builds fission-fusion societies, characterised by sex-biased dispersal and complex hierarchies. Elephant numbers have been decreasing throughout Africa due to historic culling and current poaching, as well as habitat fragmentation. This potentiates a decrease of genetic diversity, which can have negative impacts in the population. Moreover, older individuals, which act as repositories of knowledge for younger elephants, are often targeted by hunters. The removal of kin and social associates can lead to lasting behavioural changes, by diminishing strong bonds and reducing reproductive output.

We studied a South African savannah elephant population in the Greater Kruger region on the border of Mozambique to determine the genetic and relationship ties that bind individuals, and compared our results to 46 Kruger Park elephants. Our sampled elephants include 80 females and 168 males, of which 68 are juveniles and 180 are adults. We assessed the level of genetic diversity, and found no evidence of a genetic bottleneck, inbreeding or of genetic structure either within or between populations or for each sex. We also estimated the effective population size. We analysed the pairwise genetic relatedness between males in the study sample to create a genetic network, which we mapped against preliminary association data obtained from field observations of sixteen individually identified bulls. Genetic relatedness and associations were not correlated for these individuals. However, we found evidence that related bulls have at least partially overlapping home ranges. Overall, our results indicate a good genetic health for this population, conducive for the observation of male behaviours of an undisrupted population.

This pilot study offers a first genetic characterisation of a South African elephant population, and is a stepping stone towards a thorough characterisation of how genetic relatedness can explain such behaviours as males' dispersal and choice of associates.

Keywords: *Loxodonta africana*, population genetics, genetic structure, social behaviour

Index

Acknowledgments ii
Resumo iii
Abstract v
Index of Figures vii
Index of Tables viii
1. General Introduction 1
2. Genetic diversity of a growing population 4
2.1. Introduction 4
2.2. Methods 4
2.2.1. Study site and sample collection 4
2.2.2. DNA extraction and amplification 5
2.2.3. Population genetics 6
2.3. Results 6
2.3.1. Sample storage comparison 6
2.3.2. Microsatellite set validation 7
2.3.3. Population genetics 8
2.4. Discussion 11
3. Genetic differentiation of an historically fenced population 13
3.1. Introduction 13
3.2. Methods 13
3.3. Results 14
3.4. Discussion 17
4. The impact of genetic relatedness on male social associations and home range overlap 19
4.1. Introduction 19
4.2. Methods 20
4.3. Results 21
4.4. Discussion 23
5. Final Discussion 25
References 27
Supplementary Data 34

Index of Figures

Figure 1.1. Location of the APNR (bottom circle). Map by Elephants Alive 2
Figure 2.1. Electrophoresis gel for four samples in ethanol and four samples in salt solution. 7
Figure 3.1. Population structure results for individuals in the APNR and in Kruger Park..... 15
Figure 3.2. FCA analyses for individuals in the APNR and in Kruger Park.. 16
Figure 4.1. Correlation between pairwise genetic relatedness and distance..... 21
Figure 4.2. Violin plot for the correlation between social and genetic networks. 23
Supplementary Figure 4.1. Genetic network of adult males from APNR 36

Index of Tables

Table 2.1. Estimation of the effective population size using the linkage disequilibrium method.. 9

Table 2.2. Estimation of the probability of a bottleneck. 9

Table 2.3. Genetic diversity for the APNR and Kruger Park populations..... 10

Table 4.1. Pearson's r for the correlation between genetic and social networks. 22

Supplementary Table 2.1. Estimation of He, Ho, HWE, null allele frequencies, and allelic dropout and false allele rates. 34

Supplementary Table 4.1. Correlation coefficients between observed and expected values of relatedness. 36

1. General Introduction

Population genetics impact behaviour, conservation, and evolution. Genetic patterns have been shown to modulate social associations[1], dispersal[2], or mating choice[3], among others. These three biological subjects are tightly linked, especially in social species. In fact, conservation threats greatly disturb animal societies, by significantly altering social relationships[4], breeding success[5], individual survival[6], and genetic patterns[7], and thus constitute important contemporary evolutionary forces.

In social species, the removal of kin and/or social associates can result in deep social changes that may last decades and impact reproductive fitness[8,9]. However, there are many species for which it can be difficult to distinguish natural from disrupted sociality. This difficulty derives not only from the complexity of social associations[10], but also from the fact that many species are threatened, either due to climate change and its inherent consequences, or to more direct anthropogenic action, such as poaching, culling, or the destruction of natural habitats. The issue becomes particularly difficult when the population's social networks and community structure are unknown prior to a population's decline. When no basis of comparison is available to determine if a network has been affected, it is necessary to assess the genetic status of a population and search for genetic indications of a reduction in the population.

The African savannah elephant (*Loxodonta africana*) is a pertinent system to study these questions. It is a highly social species, considered by the IUCN as either Vulnerable or Endangered depending on the African region[11]. Elephant species are of great conservation concern due to human-elephant conflict and a high level of poaching. In less than a decade between the 1970s and 1980s, poaching eliminated elephant populations and reduced others by up to 90 percent[12]. What's more, the removal of individuals was selectively aimed at older animals with large tusks, which has left a significant mark on the populations, both phenotypically and socially[8,9]. Illegal hunting, which is increasing once more, habitat fragmentation, human-wildlife conflict, and a rapid decrease of living space are currently threatening savannah elephants. Indeed, wildlife is constrained to live in increasingly smaller and more isolated areas, which can impede gene flow and hasten population decline of large animals as competition grows for fewer resources[13,14].

Also of biological interest is the fission-fusion nature of elephant societies, which permit animals to react adaptively to habitat changes or threats by modifying the number and identity of individuals they associate with[15]. Fission-fusion groups are highly dynamic, merging and splitting through time and space[10]. Animals in these groups maintain stable individual relationships despite frequent changes in group composition, cohesion, duration and size, dependent on socioecological constraints[1,16]. In elephants, these groups are sex-specific, with female, kin-bonded groups also including male calves. Upon reaching adolescence, males will leave female herds and join other bulls. Both female and male groups are very dynamic and variable in terms of size and duration. Females build core groups of two to twenty adult females and their offspring. Over the course of days or weeks, core groups are divided into units as small as a single adult and her offspring, or can fuse with other core groups. Males, meanwhile, are mostly solitary, and form wide-ranging social bonds with other males, which are difficult to predict[17].

Elephants have large home ranges, the size of which is different for each population and dependent on sex, season, and habitat quality, with females showing a higher level of philopatry and site fidelity than males[18]. Elephants' use of space is highly dependent on the population, and the ecology and resources of each given region, and can undergo vast migrations in search of water and food resources[19]. However, migrations can be curbed due to presence of fences throughout Africa to limit reserves.

Many effects of population decline on elephant sociality have been described. In females, populations that suffered a greater poaching intensity tend to have more social groups composed of unrelated individuals (disrupted groups[17]). Moreover, the relationship between different disrupted

groups is more competitive than between kin-bonded groups, and females from disrupted groups show weaker social bonds, lower reproductive rates, and higher levels of stress hormones[8]. There are also cases when females are unable to form new relationships. In a park where there was a 75% population decline, 30% of groups consisted of single females[20].

Poaching has also been shown to increase male reproductive skew. By removing most of the older bulls, the remaining older males dominate reproduction due to the absence of competitors of the same age[9]. The removal of older individuals may also lower breeding age, giving males a higher chance to sire more calves than in an unpoached population[9]. Reducing the number of breeding males can augment the risk of inbreeding and genetic diversity loss, which can be exacerbated due to habitat fragmentation[17]. Seeing as social disruption can greatly impact female herds, it is vital to determine if the same can be said of males, as other negative effects may be present that have not yet been reported.

The number of elephants in South Africa is currently increasing, thanks to a fast population recovery. However, due to the elephant's long longevity (the maximum lifespan in the wild is estimated at 74 years[21]) and low reproductive rates (gestation lasts 22 months[22]), the signs of previous population shrinkage may have left a genetic and social signature still observable today. This may be particularly true in the general Kruger area, by the border of Mozambique, where culling was undertaken more extensively. Indeed, Mozambique has an estimated deficit of ~46,000 elephants and Kruger Park of 47,999 to 36,000 elephants compared to ecological benchmarks[23].

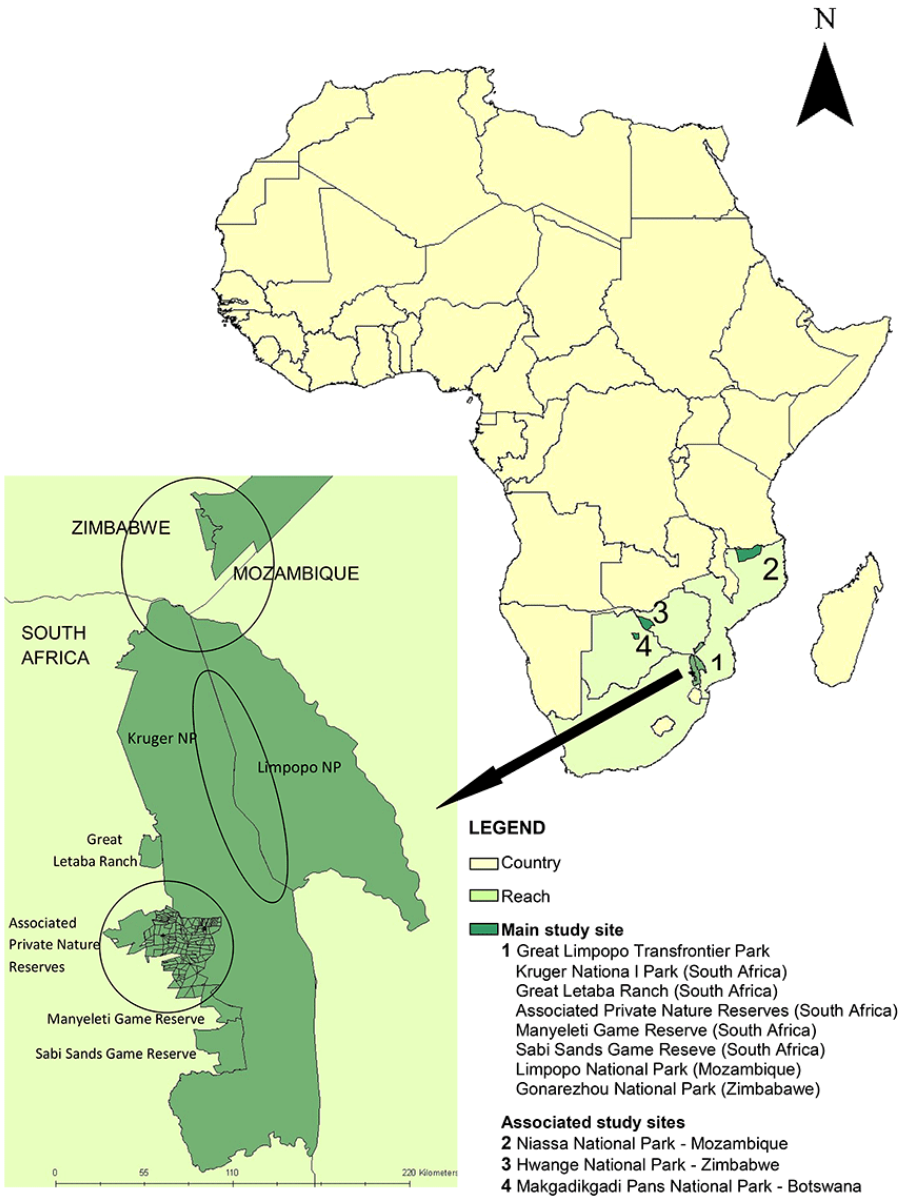


Figure 1.1. Location of the APNR (bottom circle). Map by Elephants Alive.

The Associated Private Nature Reserves (APNR) in South Africa, where this study was undertaken, are a patchwork of privately owned nature reserves that cover circa 1800 square kilometres (Fig. 1.1). These are continuous with the Kruger National Park and Mozambique's Limpopo National Park. In contrast to smaller fence reserves in South Africa, since the early 1990s no fences separate the APNR from the two National Parks, hence allowing free movement of animals between these regions. Field based research

has been conducted since 2004 on the elephants that inhabit this area. The organisation Elephants Alive has identified to date over 1500 individual elephants and, since 1998, over 50 elephants have been equipped with GPS collars that allow for tracking their movements and better understand their habitat use[24]. However, little is known about the genetic profile of the individuals or of the whole population.

Genetic tools can complement field based research by contributing important information for species conservation, such as patterns of migration[25], monitoring elusive populations[26], tracking sources of illegal animal trafficking[27], or shedding a light on the behavioural ecology of elephants that eat crops[28].

One of the most widely used tools to gather genetic data from faecal samples are microsatellite markers. Microsatellite genotyping is relatively cheap, rapid and requires only a small amount of DNA, such as that obtained from faecal matter[29]. Despite being easy to collect, it has the disadvantage of providing degraded DNA in small quantity. As such, microsatellite markers are often used instead of other genetic markers or genomic approaches, as these techniques generally require good quality DNA. Despite the high potential of next-generation sequencing to generate a large amount of data, demonstrations of its efficiency for low-cost genome-level sequencing from non-invasive low-quality samples are thus far limited in scale[30]. A recent study[31] used faecal samples to infer paternity and was only able to obtain low-coverage genomic data (0.49x) with a high degree of genotypic uncertainty. What's more, the authors highlight that current methods of pedigree reconstruction are unable to cope with the level of genotypic uncertainty in low-coverage genomic data from degraded samples. This is not an issue observed with microsatellite markers. Even though allele calling in microsatellite markers can have a high genotyping error rate, it is cheaper to produce repetitions and obtain high quality data than with sequencing based techniques[31,32].

In comparison with other markers, such as single nucleotide polymorphisms (SNPs), microsatellites are considered to be more variable, more informative, and have an approximately six times higher power for kinship analysis than SNPs[30]. Therefore, it is necessary to develop a high quantity of SNPs to match microsatellite power. For heritability estimations, hundreds to thousands of SNPs are needed to reach the same accuracy when doing pedigree-based estimations[33]. Such a quantity of SNPs is unavailable for most non-model species, an issue that is aggravated when using non-invasive sampling[30]. On the other hand, parentage detection using SNPs can be hindered by statistical problems as the lack of variability of SNPs in some populations diminishes the parent-exclusion power[34]. As such, until laboratory methodologies and post-processing techniques for efficient and cost-effective genotyping of low-quality DNA are developed, microsatellites will continue to be the marker of choice for genetic relatedness and other population genetics studies in non-model, wild species.

The main objective of this pilot study is to study the current population genetics of a population in southern Africa, as well as unveil the role of genetic relatedness in the formation and maintenance of male elephant groups. Behavioural studies require multiple years of animal observation. As such, this study has two major components: a genetic characterisation of the APNR population, and a preliminary analysis of the correlation between genetic relatedness, sociality and spatial dispersion. The latter section will be further developed in the future. We seek to respond to three main questions: what is the level of genetic diversity in the APNR? Is there genetic differentiation within and between the APNR and Kruger park populations? Do males prefer to associate and use the same home ranges as genetically related males?

We hypothesise that current and historic poaching has led to the erosion of genetic diversity of the savannah elephant and created a genetic bottleneck still visible today. We also predict that historic erection of fences between reserves has led to genetic differentiation between populations, which was further exacerbated through the action of genetic drift. Finally, we posit that, if association with kin is beneficial, social association among adult male elephants is correlated with genetic relatedness.

2. Genetic diversity of a growing population

2.1. Introduction

Several studies on the population genetics of African elephants in East Africa have been published[35,36]. However, little research has been done on elephants in southern Africa. More of these studies are important as these populations are recovering from population decreases. In fact, culling and hunting led to a bottleneck of the elephant population of Addo in South Africa[37]. Demographic bottlenecks may have a lasting social effect on elephant populations by diminishing strong social bonds and consequently decreasing individual fitness[38], and may even lead to phenotypic changes in the population. For example, selective harvesting of larger tusks has led to a decrease in tusk size[39], and to an increase of tuskless individuals[40]. It is possible, therefore, that this trait might disappear as males with smaller tusks show higher fitness. Other phenotypic and genetic characteristics may also be suffering erosion, including traits of great importance for the health of the population, namely those related to the ability to fend off disease and parasites.

It is particularly important to maintain genetic diversity as savannah elephants are classified as a vulnerable species[11]. We aim, therefore, to characterise this population genetically, with a focus on four issues: genetic variability, effective population size, occurrence of a historically recent genetic bottleneck, and presence and level of inbreeding. These subjects will serve as the first genetic characterisation of the APNR elephant population, and the information obtained in this study will aid the conservation efforts towards this population.

2.2. Methods

2.2.1. Study site and sample collection

The Associated Private Nature Reserves (APNR) in South Africa are a patchwork of privately owned nature reserves that cover circa 1800 square kilometres. These are continuous with the western boundary of the Kruger National Park and Mozambique's Limpopo National Park, and there are no fences separating the APNR from the National Parks. Field based research has been conducted since 2004 on the elephants in this area. The organisation Elephants Alive has identified to date over 1500 individual elephants[24].

A total of 360 African savannah elephant samples were collected from this study site between late October 2015 and mid-April 2017. Forty-five samples were preserved in absolute ethanol and the remaining ones in a concentrated salt solution. For 29 faecal samples there was a replicate from the same dung preserved in both salt solution and ethanol. All samples were freshly collected from elephants that were observed defecating. GPS coordinates were obtained for every sample upon collection. Other information obtained on the field upon sample collection concerned an estimation of each individual's age, whether they were seen in isolation or in a group, and group composition. IDs were given to every individual based on ear patterns and tusks observed in photographic evidence.

In addition, 46 samples, collected for a different study in the Kruger National Park, were kindly provided by Sam Wasser's team. These samples were already diluted to an average of 20 ng/ul DNA in 'Low TE' (10 mM Tris–HCl, 0.1 mM EDTA, pH 8.4).

2.2.2. DNA extraction and amplification

DNA from all salt solution and ethanol samples was extracted using the QIAamp DNA Stool Mini Kit (QIAGEN). The extractions followed the protocol suggested by QIAGEN with the following changes: all centrifugation steps were performed at 13,000 rpm, 2 ml microcentrifuge tubes instead of 1.5 ml tubes were used, and the DNA was eluted in two stages of 50 µl in a total volume of 100 µl for salt solution stored samples, or in two stages of 25 µl in a total of 50 µl for ethanol samples. The different final volumes are due to the average DNA concentration each sample type provided in the extraction test described in the next paragraph.

The same sample was extracted twice in separate extraction batches for 37 samples. Two to eight replicate samples in salt solution or ethanol for 12 known individuals were sampled at different occasions. Samples were quantified through fluorescence using FLUOstar OPTIMA, and diluted with Low TE to a concentration of approximately 20 ng/µl of DNA. To determine DNA quality for the two sample storage types, 2 µl of four ethanol and four salt solution samples were loaded onto a 1% agarose gel with ethidium bromide (Fig. 2.1), along with a 25 ng/µl lambda standard. The gel was run for 45 minutes at 95V.

Subsequently, preliminary genotyping was performed to assess microsatellite amplification. This test included 26 salt-stored samples and 19 ethanol-stored samples. They were genotyped for five autosomal species-specific markers, namely FH19, FH71[41], FH126, FH127 and FH153[42], and three sexing markers, PLP1, SRY1 and AMELY2[43]. The last two are Y-linked markers with the aim to maximize the chance of correct sexing, as the dropout of Y-linked markers could lead to a negative bias in the identification of males.

Each singleplex PCR reaction contained 1 µl of primer mix (containing fluorescently labelled forward primer and unlabelled reverse primer) at 0.2 µM, 1 µl QIAGEN Multiplex PCR Master Mix (QIAGEN), and either 1 µl or 2 µl air-dried DNA. 1 µl was used if the sample's concentration ranged between 10 and 20ng/µl and 2 µl if the DNA's concentration was below 10ng/µl. PCR amplification was performed using a DNA Engine Tetrad Thermal Cycler (MJ Research). The PCR temperature protocol consisted of an initial denaturation at 94°C for 15 min, followed by 45 cycles of 30 seconds at 94°C, 90 seconds at 56°C, and 90 seconds at 72°C, followed by a final extension at 72°C for 10 min. PCR products were then diluted 1/800 with double-distilled water and resolved using an ABI 3730 48-well capillary DNA Analyser (Applied Biosystems, California, USA). Allele scoring was performed using GeneMapper 3.7 (Applied Biosystems, California, USA). Negative controls were included in the DNA extractions and PCRs to control for contamination.

Based on 26 unrelated individuals, expected and observed heterozygosities were calculated and null alleles frequencies estimated for each marker using CERVUS 3.0.6 software[44]. Deviation from Hardy-Weinberg equilibrium and linkage disequilibrium were assessed using GENEPOP 4.2[45]. The possibility of scoring errors due to stuttering, large allele dropout and null alleles was assessed in Micro-Checker 2.2.3[46].

Once the consistency of this protocol was assured, 18 autosomal microsatellite markers were chosen to genotype all samples. The species-specific markers amplified were FH1, FH19, FH39, FH40, FH48, FH60, FH67, FH71, FH94, FH103[41], LA5, LA6[47], LaT06, LaT08, LaT13, LaT18, LaT24, and LaT25[48]. The reverse primers contained a pigtail with the motif "GTGTCTT" at the 5' end. These markers were chosen as they were previously used to genotype the 46 samples lent by Sam Wasser's team, thus allowing more individuals to be analysed and different sites to be compared. The markers were divided into three multiplexes. To verify the obtained genotypes, all samples were genotyped at least twice. In the presence of incongruous results or of homozygotes, the samples were repeated at least twice more.

Once all genetic data was collected, it was analysed for quality. The analysis was conducted for all samples, as well as for samples in salt solution and ethanol separately to determine the differences between outputs for each storage method. To control for any scoring discrepancies, Micro-Checker was

used. All repeats for the same sample were compared both manually and using the software GIMLET 1.3.3[49], on which consensus genotypes for each sample were built. A consensus genotype was built when the same allele amplified at least twice. GIMLET also permitted the assessment of the level of allelic dropout, false alleles, and five types of genotyping error through a pairwise comparison of all genotypes for the same sample. These calculations were based on 53 randomly selected samples. The data was also manually checked.

Probability of identity (PID) was assessed using CERVUS to ascertain that different samples for the same known individual had the same genotype, and to identify samples that belonged to the same unnamed individuals. Two samples were accepted as belonging to the same elephant when their respective genotypes had up to two loci mismatches.

As above, deviation from Hardy-Weinberg and linkage disequilibrium were assessed using GENEPOP, and expected and observed heterozygosities were calculated and null alleles frequencies estimated in CERVUS, based on 20 unrelated individuals. The individuals were selected by calculating the pairwise relatedness between all individuals on the software ML-Relate[50]. The 20 chosen individuals correspond to those with the lowest pairwise genetic relatedness.

2.2.3. Population genetics

The effective population size (Ne) was estimated in NeEstimator 2.01[51] under the software's default parameters for the linkage disequilibrium (LD) method, assuming a random mating model. The Ne was calculated for critical values (Pcrit) of 0.01, 0.02, and 0.05.

Evidence for the existence of a bottleneck was investigated in Bottleneck 1.2.02[52] using three different statistical tests: the sign test, the Wilcoxon test, and the mode-shift test. The first two test types were run assuming either the infinite allele (I.A.M.), two-phase (T.P.M.), or stepwise mutation (S.M.M.) models. The sign and Wilcoxon tests detect whether a significant number of loci exhibit heterozygosity excess, as expected in a population undergoing a recent bottleneck. The mode-shift test assesses the distribution of allele frequencies, which is expected to be L-shaped under mutation-drift equilibrium[53]. All estimations are based on 10,000 replications. The T.P.M. was run with a variance of 12 and a probability of 95%, and with a probability of 78%[52,54].

Both the bottleneck and effective population size tests were performed with all APNR and Kruger individuals pooled together, as using a small sample size when testing for these events is not recommended[55]. Furthermore, individuals collared in the APNR have been observed in the Kruger National Park, thus indicating that elephants in both these sites might belong to the same population (also supported by the results in Chapter 3).

The level of inbreeding (Fis) was assessed in Genetix 4.05.2[56] based on 1,000 permutations. The overall Fis and the Fis for each microsatellite marker were calculated. HP-Rare[57] was used to determine the allelic richness and the existence of private alleles in the APNR and Kruger populations. This was done using a rarefaction method to account for variance in samples size among regions.

2.3. Results

2.3.1. Sample storage comparison

The samples stored in salt solution had an average concentration of 33.492 ng DNA/µl, whilst ethanol samples had an average concentration of 7.888 ng DNA/µl.

When run in an electrophoresis gel, the samples stored in ethanol showed clearer smears than those stored in salt, with the latter showing stronger single bands (Fig. 2.1). As such, it is expected that samples stored in salt have the least degraded DNA.

The average amplification rate differed between storage types. For ethanol stored samples, 98.89% of the autosomal loci amplified for all individuals. The percentage was 94.17% for salt stored samples. Conversely, we could assign sex with complete certainty to 92.31% of the salt samples and 89.47% of ethanol samples (i.e. females only amplify the X-linked locus PLP1, and males amplified both the X and the two Y-linked loci). All sex-typing results matched field observations when these were available.

Ethanol stored samples showed slightly higher values of allelic dropout, false alleles, and types 1 and 3 errors than those in salt solution. However, the differences between storage methods were always under 2%.

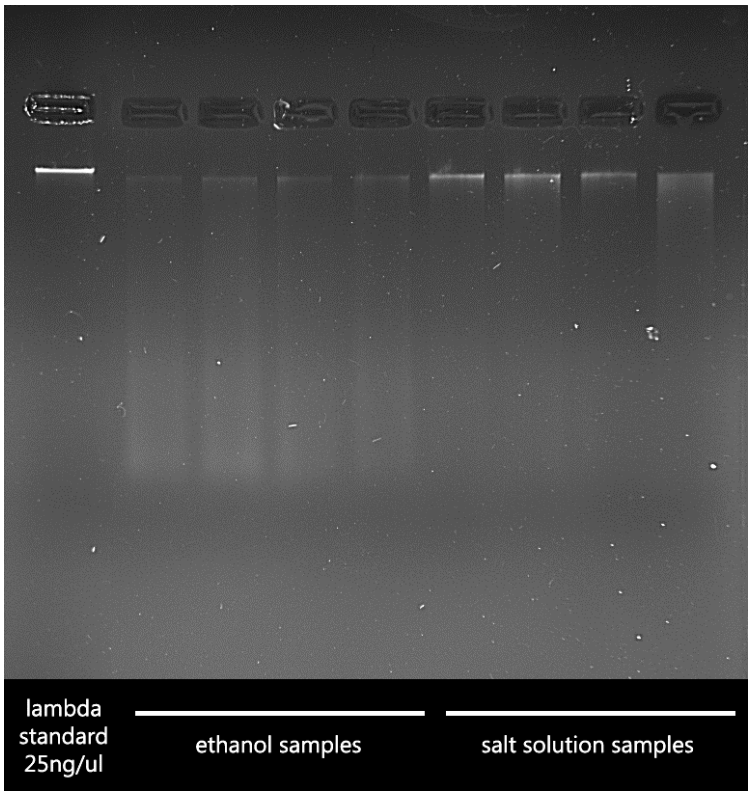


Figure 2.1. Electrophoresis gel for four samples in ethanol and four samples in salt solution.

2.3.2. Microsatellite set validation

Two different subsets of unrelated samples were used to determine whether the chosen microsatellite markers were suitable. A slight deviation from Hardy-Weinberg equilibrium was detected for three markers and deviation of linkage equilibrium (LE) was observed between two markers, after applying a correction for multiple testing[58], but each deviation was only present in one of the sample subsets. Therefore, the lack of consistent deviations from HWE and LE indicate that the observed deviations are intrinsic to the samples themselves, and not due to specific problems with the markers, and thus support the suitability of the chosen marker set (Supplementary Table 2.1).

The rates of false alleles and allelic dropout were highly variable within the market set. However, even though some markers had a comparatively high rate of false alleles and/or dropout, the percentage of these remained below 5% regardless of the calculation method (across all loci, samples, and PCRs).

The quantity of Type 1 to 5 errors was much smaller, ranging from 0 to 0.9%. It is possible that the quantity of false alleles was overestimated due to the parameters chosen for this evaluation. A consensus genotype was only built when an allele was observable at least twice. As such, it is possible that a sample that amplified as a homozygote several times, and only once as a heterozygote is classified as a homozygote even though it is actually a heterozygote. This would lead to the classification of this "extra" allele as false, and boost the false alleles rate. According to the Oosterhout's estimator, there is also a possibility of the presence of null alleles in three of the 18 microsatellites (FH71, LaT06, and LaT25). For these, the likelihood of null alleles was higher than 0.05 (Supplementary Table 2.1).

Taking all these observations into consideration, we concluded that the chosen marker panel provided a reliable genotyping of the study samples. From the genotyping of all available samples we also estimated the number of individuals. In total, from 410 samples, we genotyped 294 different individuals, including those of the Kruger region. Of these, 98 were females and 196 were males.

2.3.3. Population genetics

The impacts of recent history on current demography was investigated. We began by estimating Ne. The choice of the best Pcrit for this estimation should be based on sample size[59]. In our case, only alleles with frequencies higher than 0.02 should be considered[60]. From the two alternatives to estimate confidence intervals, the 'jackknife on loci' method is considered the most reliable[51]. Under these considerations, the Ne was estimated to be of 394.4 individuals, with the 95% confidence interval boundaries being 271.1 and 657.5 (Table 2.1).

There is no evidence of a recent bottleneck in this population, as the p-values obtained in the T.P.M. (95%) and S.M.M. for the one-tailed Wilcoxon test of heterozygosity excess indicate that it is not possible to reject the null hypothesis of mutation-drift equilibrium (Table 2.2). The significant global sign tests were clearly due to heterozygosity deficiency not excess, and the mode-shift test found a normal L-shaped distribution of allele frequencies, as expected under mutation-drift equilibrium. The TPM at 78% probability did not reject the null hypothesis in both the Wilcoxon and sign tests. The I.A.M. tests pointed in the opposite direction, but this model is generally not accurate for microsatellite data.

The overall values of Fis for both the APNR and Kruger Park populations were significant and slightly above zero, indicating that some inbreeding may be present in the population (Table 2.3), but the values were nonetheless low. The allelic richness was similar between the two populations, but slightly higher for the APNR. The same trend was observed for the private alleles proportions.

Table 2.1. Estimation of the effective population size using the linkage disequilibrium method for different Pcrit values. CI: confidence interval.

Lowest allele frequency used	**0.05**	**0.02**	**0.01**	**0**
Harmonic Mean Sample Size	287.5	287.9	287.4	287.0
Estimated Ne	330.5	394.4	364.6	402.9
Parametric 95% CI for Ne	260.4 438.5	318.1 508.1	306.8 443.3	343.6 482.1
Jackknife on Loci 95% CI for Ne	212.8 616.9	271.1 657.5	270.9 529.4	265.3 732.9

Table 2.2. Estimation of the probability of a recent bottleneck having occurred in the study population based on 10,000 replicates. All tests were conducted under the null hypothesis of mutation-drift equilibrium. I.A.M.: infinite allele model. T.P.M.: two-phase model. S.M.M.: stepwise mutation model. h(e): heterozygosity excess. h(d): heterozygosity deficiency. P: probability.

Model	**Sign test**				**Wilcoxon test**		
	Expected n loci with h(e)	N loci with h(d)	N loci with h(e)	P	P. 1-tail for h(d)	P. 1-tail for h(e)	P. 2-tails for h(e) or (d)
I.A.M.	10.31	2	16	0.00432	0.99987	0.00017	0.00034
T.P.M. 95% p	10.63	15	3	0.00029	0.00117	0.99903	0.00233
T.P.M. 78% p	10.55	10	8	0.16303	0.07702	0.92924	0.15405
S.M.M.	10.64	16	2	0.00004	0.00005	0.99996	0.00010

Table 2.3. Genetic diversity for the APNR and Kruger Park populations. *: significant values for alpha = 0.05.

	APNR			Kruger		
Locus	Fis	Allellic Richness	Private Alleles	Fis	Allellic Richness	Private Alleles
Overall	0.03018*	5.47	0.96	0.05388*	5.28	0.77
FH1	0.01334	4.6978	0.239	-0.00712	4.7016	0.2428
FH19	0.06994*	5.9251	0.5078	0.02641	6.9385	1.5212
FH39	-0.04881	6.9076	1.1807	0.01915	7.7376	2.0107
FH40	-0.07039	4.3133	0.4635	0.04218	3.9205	0.0707
FH48	-0.01510	5.2967	0.8988	-0.02118	5.4708	1.073
FH60	-0.03909	2.9029	0.2559	0.05821	2.7792	0.1322
FH67	0.06186*	6.2514	1.5956	0.06427	4.8806	0.2249
FH71	0.11350*	2.9532	0.0434	0.08924	2.9545	0.0447
FH94	0.04184	4.7226	0.7886	-0.12017	3.9495	0.0156
FH103	0.03996	4.3818	0.7402	-0.12749	4.1557	0.5142
LA5	-0.01844	4.1347	0.9276	-0.01916	3.7049	0.4979
LA6	-0.06158	2.5177	0.1929	0.01296	3.5619	1.2371
LaT06	0.17058*	6.8251	2.5216	0.27940*	5.4451	1.1416
LaT08	-0.02167	9.5215	1.6002	-0.01038	9.9612	2.0398
LaT13	0.07319*	6.7363	1.5269	-0.01026	6.0203	0.8109
LaT18	-0.00770	6.3161	1.1572	0.39254*	5.7566	0.5977
LaT24	0.02024	8.4473	2.0484	0.20635*	7.1249	0.7261
LaT25	0.17027*	5.6901	0.6489	-0.02467	6.0363	0.9951

2.4. Discussion

Despite the historical and current evidence of anthropogenic action on the elephants in the study area and their habitat, we were unable to detect evidence for a genetic bottleneck, severely reduced genetic variation, or substantial inbreeding. The estimated effective population size was also high, exceeding the number of sampled individuals. The heterozygosity deficiency for the observed allelic diversity suggests that the population might be in fact undergoing an expansion. Although an increase of the number of elephants has been observed in the field (Mumby *pers. comm.*), additional genetic tests are needed to ascertain if a genetic diversity increase has indeed occurred.

The average Ne/Nc (effective population size/number of censused individuals) ratio for wild populations is about 0.1[60]. The last census of elephants in the APNR was conducted in 2012, when it was estimated that circa 1,500 elephants occupied the area, and the highest number of individuals observed in the region over a one year period was of about 2,000 in 2006. As field observations point towards a gradual increase in population numbers, it is possible for elephant numbers to be proximate to those of 2006. If that is the case, the Ne/Nc ratio herein observed is about 0.2, higher than the average of wild populations.

A high Ne suggests that there is a slow loss of genetic diversity. This is further supported by the fact that no bottleneck or substantial inbreeding was detected. Elephants' breeding group structure may contribute to the slower action of genetic drift. The division of the population into breeding groups may prevent genetic diversity loss, as it is usually accompanied by sex-specific dispersal[61], which will add diversity to each gene pool. Also, a small number of migrants may suffice to mask the genetic signature of bottlenecks[62,63]. Documented immigration from Mozambique may have increased the number of rare alleles in this population without a corresponding impact on heterozygosity[37].

A study focused on South African Kruger and Addo elephant populations compared the genetic diversity of these populations against that of a Ugandan population, by looking at the allelic diversity of four loci shared between studies[37]. They observed that South African populations are less genetically diverse than populations in Uganda. Unfortunately, we were not able to perform such a comparison as no studies other than those where the markers were first described could be found that reported allelic diversity for the same loci we used[41,47,48]. Of the eighteen markers used, only three had a lower allele number in our study (FH60, FH71[41], and LA6[47]). However, a limited number of individuals was used in the studies describing the markers (10[47] to 23[41] elephants), which may render these studies inadequate for an allelic diversity comparison.

A recent study[64] has measured the levels of overall heterozygosity in various regions of southern Africa. According to the values therein reported, the APNR population, with an average observed heterozygosity (Ho) of 0.614 for all loci, falls within the range of Ho values for these populations (from ca. 0.4 to ca. 0.76). This same study found that, similarly to the APNR, the Kruger Park population had a Ho of ca. 0.6. This is not the case when comparing the APNR population against eastern Africa. Indeed, six of seven populations throughout Kenya showed a higher Ho than the APNR (from 0.655 to 0.690)[65], as did five regions in southern Kenya and northern Tanzania (0.637 to 0.751)[66]. As such, the APNR population has the average Ho of other populations in southern Africa, but a lower Ho in comparison to populations in eastern Africa. This is not surprising, as savannah elephants in southern Africa have a lower genetic diversity than elephants in other regions, especially when forest elephants are also considered[67].

Overall, it appears that the APNR and Kruger populations are genetically healthy, but this does not signify that little conservation action is necessary for this population. With elephant numbers dwindling across Africa, it is of the utmost importance to maintain the social and genetic diversity of the APNR population to safeguard the future of the species. In fact, the lack of genetic differences between social groups is conducive to a higher risk of losing alleles to an increase of genetic drift[68]. It is necessary, therefore, to ensure that social groups are maintained. Furthermore, given the life-span of

elephants, it is possible that it is too soon to observe any significant loss of genetic diversity. As such, it would be of interest to genetically reassess the genetic diversity of South African populations in the next decades to better inform conservation policies.

3. Genetic differentiation of an historically fenced population

3.1. Introduction

Studies of genetic structure find the level of differentiation between groups of individuals of the same species in a geographical area by assessing the variation of allele frequencies between individuals, and grouping those that share similar allele frequencies into separate clusters. Genetic structure is not always replicated on the geographical distance between individuals[69]. A population where no physical obstacles are identifiable may nonetheless be genetically structured due to the presence of unidentified barriers to gene flow. On the other hand, individuals that inhabit distant locations may belong to the same genetic cluster, as migrants can be sampled away from their natal cluster[69]. Furthermore, dispersal provides a higher chance of breeding with distant individuals, and first generation descendants of migrants may be assigned to the migrant parent cluster[70].

Elephants exhibit seasonal migrations, following well-established routes[71], and breed throughout the year. It is expected for female herds to include offspring of distant individuals that follow at least part of the same migratory routes and/or visit the same sources of water. However, fences and human occupation can alter foraging and migratory paths, which may result in an absence of gene flow between populations[72,73] and intensify the effects of genetic drift.

Multiple studies have attempted to decipher the overall genetic structure of elephants. Most of these are continental phylogeographic analysis with few sampled individuals per location, hence masking differences at a smaller scale[42]. Indeed, phylogeographic studies have only been able to determine a low differentiation between savannah elephants from elephants in other African areas[74]. However, it should be noted that no phylogeographic studies of the savannah elephant have been done without including forest elephants, which can affect any conclusions concerning intraspecific diversity[75]. Other studies were conducted using a small number of microsatellite markers, which may not be sufficient to ascertain fine scale structure[76]. When over ten microsatellite markers were employed, only the genetic structure of female groups of a specific region was analysed[61], or the study was performed on distant populations, which may have different behavioural profiles from elephants in southern Africa[66,77]. What's more, the fine-scale genetic structure of elephants in South Africa remains largely unknown despite the importance of this information for species management.

Studying genetic structure allows us to infer the impact of past and current anthropogenic action in South Africa and in Mozambique as driving evolutionary forces. Moreover, the identification of genetic clusters can aid such conservation decisions as translocations and reintroductions to locations of low elephant density where human-elephant conflict has been alleviated enough for the species to persist.

3.2. Methods

Evidence of genetic structure was investigated using the programs GENELAND 4.0.8[78], STRUCTURE 2.3.4[79], and TESS 2.3.1[80] to determine the most likely number of genetic clusters (K). The combination of these softwares enables the comparison between aspatial (as in STRUCTURE) and spatial algorithms (as in the other two software). To maximise comparative power, the same parameters were chosen whenever possible, and K was tested between 1 and 8. The analyses followed the recommendations of Basto *et al.*[81], with a few alterations. In GENELAND, the allele frequency models were employed with the spatial and null allele models, the coordinate uncertainty was set at 0.01, K was inferred from 15 independent runs, and the correlated model was run with either the default medium-

sized differentiation beta prior parameters (2,20), an uninformative differentiation prior (1,1), or a low differentiation prior (1,100)[82]. Three different beta prior parameters were used because, despite being more powerful in the detection of subtle differentiation, the correlated model can also be more error prone due to a higher sensitivity to departures from model assumptions and predisposition to algorithm instabilities[83]. In STRUCTURE, 500,000 iterations were performed in each run. To obtain representative figures for the likeliest K, STRUCTURE HARVESTER[84] results were processed in CLUMPP[85] and distruct[86]. If the inferred K was one, the results for K = 2 were also processed.

A factorial correspondence analysis (FCA) was also performed in Genetix 4.05.2[56] to look into individual variability and population structure.

As elephants exhibit male dispersal, whilst females aggregate in family groups, all analyses were run for each sex separately, as well as for all individuals pooled together. Finally, pairwise differentiation (Fst) between the APNR and Kruger Park populations was estimated in GENEPOP 4.2[45].

3.3. Results

Most models used herein suggest the presence of a single genetic cluster for each sex and for all samples (Fig. 3.1). Neither the admixture model in STRUCTURE, the uncorrelated model in GENELAND, nor the non-admixture and CAR models in TESS assigned individuals to distinct genetic clusters. On the other hand, the correlated model in GENELAND was consistently unable to find a stable K, regardless of the beta prior used. The model was unable to converge on a specific K for any of the three sets of individuals under the three different sets of beta priors. No stable K was also found when using the BYM admixture model in TESS for any of the sets of individuals. Furthermore, when separate runs of the same model yielded the same value of K, different individuals were grouped together.

The factorial correspondence analyses showed that variation among individuals was similarly explained by both axes (Fig. 3.2). Two unnamed males from the APNR were very different from other males, whilst the variation between females was homogeneous. When considering all individuals, these same two males were the most different ones from the remaining cluster. There was no separation of individuals into clear groups. The overall value of Fst was 0.0038, further corroborating the little differentiation within and between populations.

Figure 3.1 Population structure results for individuals in the APNR and in Kruger Park for a) all females using the software STRUCTURE, b) all individuals using the software TESS under the no-admixture model, c) all individuals using the software GENELAND under the uncorrelated model. The dots in b) and c) correspond to the sampling location of each individual.

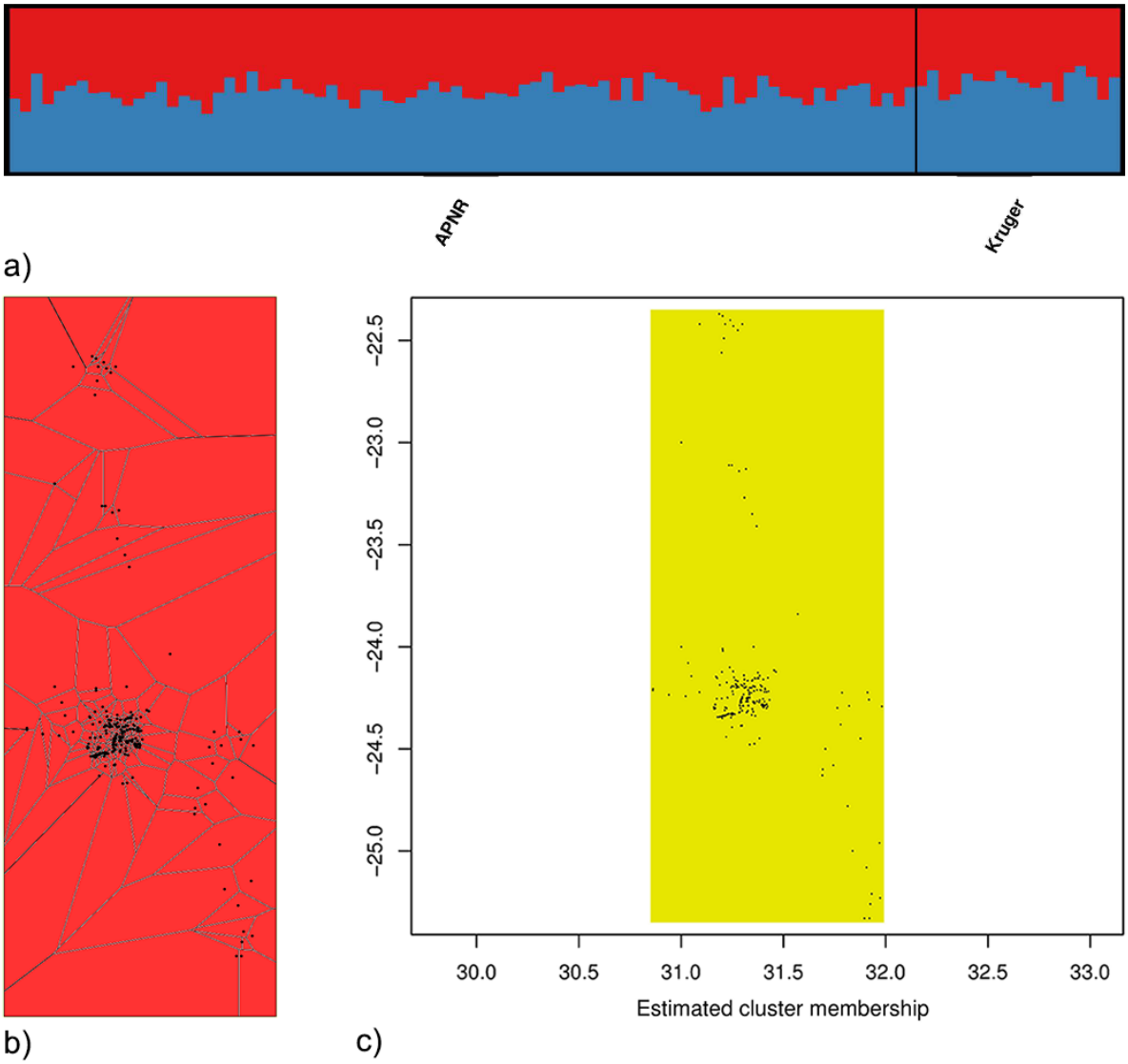
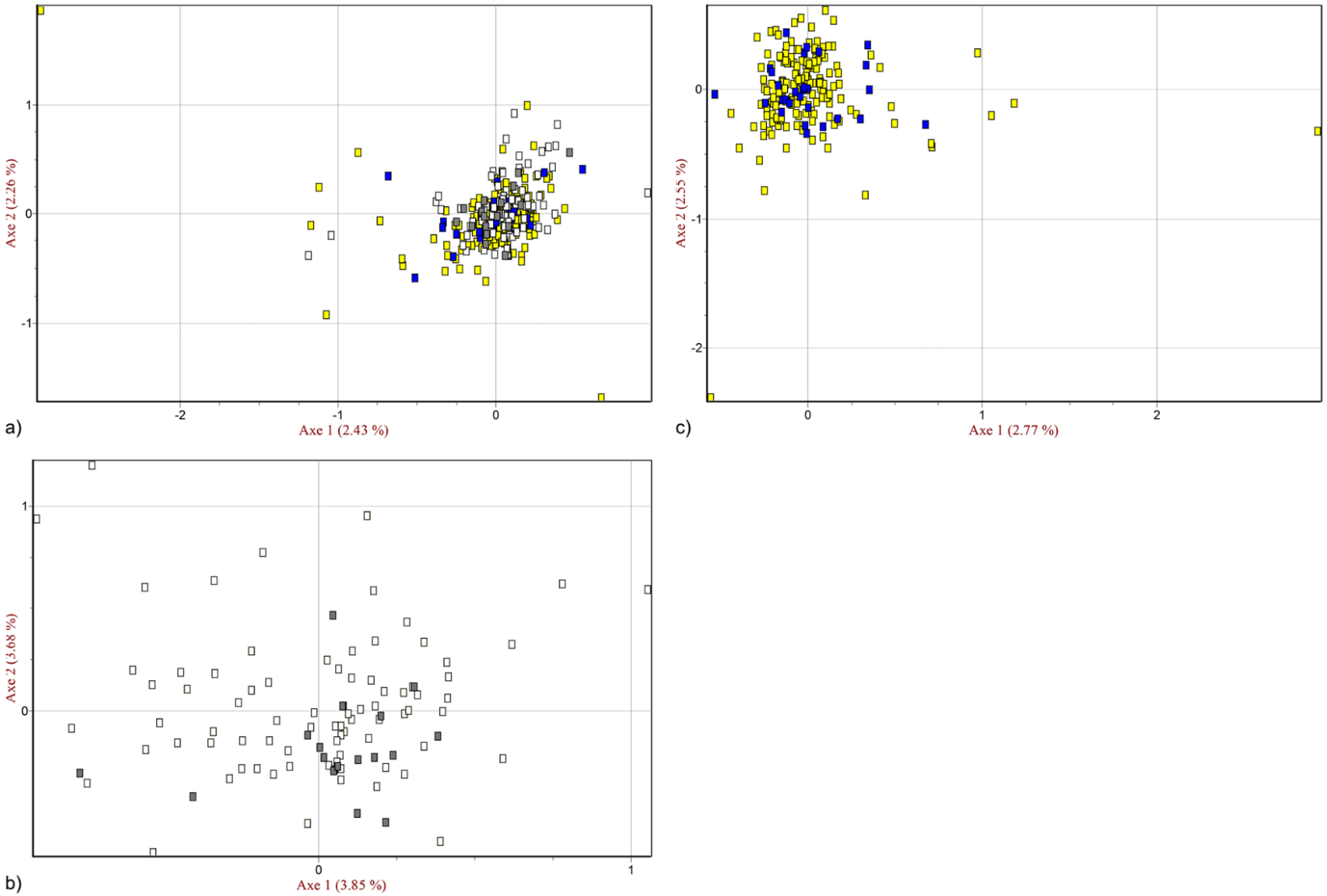


Figure 3.2. FCA analyses for individuals in the APNR and in Kruger Park for a) all individuals, b) all females, and c) all males. Yellow corresponds to males from the APNR, blue to males from Kruger, white to females from the APNR, and grey to females from Kruger.



3.4. Discussion

There is no apparent evidence of genetic structure either within or between the APNR and Kruger populations, as most models found a single genetic cluster. When that was not the case, it is likely that the models yielded false positives or could not converge. The correlated model in GENELAND takes into account that allele frequencies may be correlated across spatially proximate populations, rendering it theoretically more suitable to search for low differentiation resulting from recent ecological events. As such, it may be able to detect variation at fine spatial scales or within mobile species[82]. However, this model is also more sensitive to deviations from model assumptions and more prone to algorithm instabilities[83], and can detect extra clusters that may not reflect true genetic structure. Likewise, the admixture models in TESS have been shown to frequently infer spurious clusters, which vary greatly both in number and in spatial features depending on slight changes of the parameters used[87]. It could be possible that these more sensitive models were able to detect microstructure or family groups that the other models failed to identify. This is nonetheless unlikely as the clusters' composition differed greatly between independent replicate runs.

A small number of individuals were shown to have a more distinct genotype. These might be migrant elephants, which settled in the northwest area of South Africa due to a lower amount of poaching. In fact, the nature reserves in South Africa are more protected from poaching and culling than other regions in southern Africa, where these practices are still prevalent. Although hunting is still carried out in areas like the APNR, regulations are enforced that only allow a set quantity of individuals to be hunted per year. The presence of individuals with GPS collars is also a deterrent against poaching, as the death of such individuals is detected more rapidly and can result in the apprehension of poachers. Despite being illegal in many countries, ivory trade has already led to the disappearance of big tuskers in elephant populations and to an increase of tuskless and smaller-tusked individuals throughout Africa[40,88,89] as this is a heritable trait that has been selected against. The Greater Kruger area is currently considered one of the strongholds of big tuskers[90]. The overall elephant population in this area is granted a higher degree of protection than populations in other areas of Africa, hence making it appealing for animals.

It may be difficult, however, to detect migrants from proximate countries. The seeming inexistence of population structure over such a large area indicates a certain homogeneity of genotypes, which may result from a high level of admixture between groups of elephants. Despite the erection of fences throughout Africa, the border between Kruger National and the Limpopo National Park is porous. Free movement associated with migrations and large home ranges increase the chance of reproduction between individuals that originate from distant places, hence augmenting the genetic admixture. This pattern is not observed in populations with historical habitat and population fragmentation, which left a clear mark in the genetic structure of populations[65].

Previous studies have shown different genetic structure patterns in elephants for mitochondrial and microsatellite DNA[91], with the former allowing for a clearer identification of genetic clusters. Underlying such a striking difference may be a combination of male-biased dispersal, higher site fidelity from females, and inbreeding avoidance[92]. Male-biased gene flow could result in the homogenization of nuclear alleles, whilst maternally inherited mitochondrial DNA variants would be restricted to individual areas[93]. Male-mediated gene flow has been described in other populations of elephants[66,93], as well as in other species, such as Australian green turtles (*Chelonia mydas*)[94] and Australian red kangaroos (*Macropus rufus*)[95]. Furthermore, even though no differentiation was found when using neutral loci, it is possible for it to exist on loci under natural selection[96]. For example, a study of Atlantic cod could find little genetic differentiation between populations at microsatellite loci (Fst = 0.003), but was able to find great differentiation at a locus that was under natural selection[97]. It would be interesting to test this for the African elephant, starting, for example, with genomic areas associated with tusk size.

Genetic structure can also be influenced by other factors, such as population density and demography. An analysis of fine-scale genetic structure of red deer (*Cervus elaphus*) over a 24-year

period showed that genetic differences between herds declined from 4% to 1% after a decrease in hunting. Not only did the end of culling lead to a substantial increase of breeding females, but it also resulted in more evenly distributed mating opportunities across the population[98].

Thus, although we were unable to detect structure in this population, it does not necessarily follow that it does not exist. It may be that male dispersal guarantees a high level of gene flow, whilst group differences persist in mitochondrial DNA and even, to a lesser degree, in the X chromosome. Accordingly, previous studies have detected mitochondrial and nuclear DNA sequences that indicate the existence of genetically differentiated elephant groups[76], but diagnostic sequences at a smaller regional level have been produced for very few regions[66]. Current next-generation sequencing, paired with a thorough sampling at a continental level, could permit us to obtain a mitogenomic profile of elephants from the various areas of Africa. Such a profile, combined with diagnostic nuclear DNA sequences, could be a useful tool to estimate the geographical origin of poached ivory and identify which regions are in more pressing need of protection[27,67]. These findings could also be incorporated into future management of the species, such as translocations and the establishment of connectivity paths between differentiated populations, to help avoid the disappearance of unique haplotypes.

4. The impact of genetic relatedness on male social associations and home range overlap

4.1. Introduction

Elephants are among the most social species. They show an unusually high degree of empathy[99], can recognise around 100 other individuals based on vocal communication[100], and are concerned over distressed or deceased individuals, regardless of whether they are kin[101]. Furthermore, elephant societies have been shown to be adaptive to changes in the environment[102]. As such, it is vital to understand the underpinnings of male sociality, especially as the removal of key males by poachers or trophy hunters could destabilise social cohesion[103].

Until recently, it was thought that adult males lead a mostly solitary life with very few and random associations with other males[17], but studies now show that male elephants are more social than previously considered, particularly during non-musth periods[104]. Male groups include bulls of various ages, from adolescence to adulthood, with larger groups including a higher number of young bulls[105]. Just as for females[100], the older males are a repository of knowledge for the younger bulls[103]. However, little is known about what ties male elephants together.

Kin recognition is well documented in female elephants[15,106]. It has been hypothesized to influence female reproductive success, with females with fewer kin having a lower reproductive rate[106]. The benefits of male association, on the other hand, are mostly speculative. If kin-bonds are important for males, it may be that related males facilitate each other's reproductive success in cooperative rather than competitive interactions as observed in peacocks (*Pavo cristatus*)[107], or that kinship diminishes aggressive behaviours as documented in western gorillas (*Gorilla gorilla*)[108]. However, most studies on bull sociality focus on other factors that may impact male bonding, such as musth state (a period of two to three months in a year when males seek estrous females and show aggression towards other males), season, age, and geographic area[103,104,109]. Many of these works have conflicting results. Thus far, the only parameter that has been agreed on by multiple studies to have a marked effect on group dynamics is age, with older bulls being one of the preferred associations[103,105].

The genetic ties that connect males have only been explored in elephant populations in Kenya. One study showed that male association was weakly correlated with pairwise genetic relatedness as it was preferable to associate with males of the same age[103], whilst another found no evidence for relatedness between males observed together[77].

The use of space by elephants has been shown to differ depending on landscape heterogeneity and the availability of natural water sources[110], but each group's spatial dispersion has not been tested against genetic relatedness. Bulls are not generally territorial, but can avoid other individuals' home ranges during musth[111]. An assessment of whether genetically related males are found in the same home range more often than expected by chance can also aid in the determination of whether kinship is an important parameter in the formation and maintenance of groups.

Herein, we present a pilot study that aims to begin studying the impact of kinship on male associations for a savannah elephant population.

4.2. Methods

4.2.1. Genetic relatedness and geographical distance correlation

The R package related[112] was used to choose the best estimator for the calculation of pairwise relatedness. The estimations were run twice for 100 and 500 simulated relationships separately, based on the allele frequencies of all elephant samples. The estimator with the highest correlation coefficient between expected and observed values of relatedness was chosen.

A first approach at determining whether genetic relatedness impacts individuals' home ranges was undertaken by assessing whether there is a correlation between pairwise genetic relatedness and the spatial distance between samples in SPAGeDi v.1.5[113]. Genetic relatedness was plotted using the Queller and Goodnight relatedness estimator[114] over 20,000 permutations between all individuals in each test. The same allele frequency values were used for each test, regardless of sample number. Seven distance classes were chosen to perform the analysis: the classes between 0 and 0.5 km represent individuals within a group, those between 0.5 km and 15 km represent the distances where elephants can use seismic communication[115], and distances larger than 15 km are those where elephants are unable to communicate. The analysis was run for adults of each sex separately for all observations. The individuals observed multiple times were included with only one set of randomly chosen GPS coordinates to not bias results.

4.2.2. Genetic and social network correlation

The values of pairwise relatedness between all individuals was computed in COANCESTRY v.1.0.1.7[116] using 100 reference individuals and 20,000 bootstraps. The Queller and Goodnight relatedness estimator was chosen to perform this calculation. Using these results, a genetic network for all APNR males 20-years-old and over was constructed in the software gephi v.0.9.1[117]. This age cut-off was chosen as focal bulls in social networks are estimated to be as young as 21-years-old[104].

Three different social networks were built, taking into account observations of (1) all males, (2) males in male-only groups, and (3) males when not in musth. Only males observed at least eight times and that had an ID assigned to them in the genetic relatedness data were included. ID assignment was necessary to ensure overlap between genetic and social databases. As such, the first network assessed 16 males, the second 13, and the third 15. The association between individuals was calculated using the half-weight index, which is considered to be less biased when not all individuals were identified in a sighting[118].

The correlation coefficient between social and genetic networks was calculated using the 'gcor()' function in R. The original datasets were permuted 2,000 times to create a 'null model', wherein each individual was observed the same number of times and on the same day but associated randomly. The data structure was controlled by making 1,000 swaps of observations for each day. The correlation coefficient was calculated for each permuted dataset. A distribution was built for the permuted correlation coefficients to determine whether the results were obtained by chance. Results are expected to have arisen by chance when the observed coefficient falls within the middle 95% of the distribution. When that is not the case, there is a departure from what is expected if associations were random.

4.3. Results

4.3.1. Genetic and geographical distance correlation

The tested estimators had a similar performance when assessing their capacity to distinguish between four types of relationship (unrelated, parent-offspring, half-sibling, and full sibling; Supplementary Table 4.1). As the Queller and Goodnight estimator had a slightly higher correlation coefficient than the other estimators, it was the one chosen for all the subsequent genetic relatedness analyses.

There is evidence for a higher pairwise genetic relatedness than expected by chance between male elephant samples within 100 meters of distance, similarly to what is observed in females. This higher relatedness is lost at greater distances, when animals are no longer part of the same group (over 500 meters of distance between one another). This pattern is observed both for all individuals and for only adults of each sex separately (Fig. 4.1).

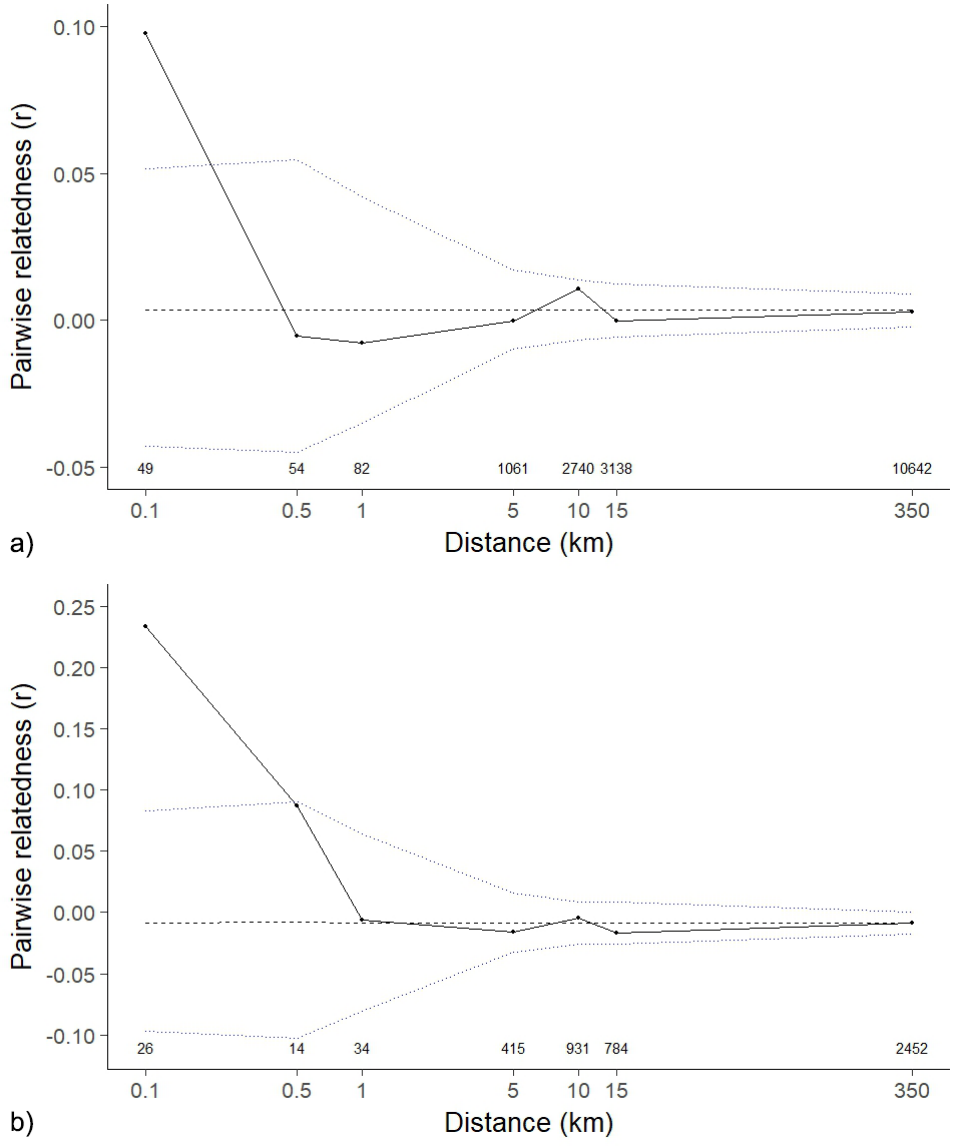


Figure 4.1. Correlation between pairwise genetic relatedness and geographical distance for all samples of a) males and b) females, with spatial coordinates, both from Kruger and APNR. Both figures include young adults and adult individuals, i.e. all females are over 10 years old and all males are over 15 years old. The dotted lines correspond to the 95% confidence

intervals, the dashed line to the average value after permutation, and the full line the average relatedness values within each distance class. The numbers above the distances indicate the number of pairwise comparisons for the respective distance class.

4.3.2. Genetic and social network correlation

We built a genetic network with all adult males that were sampled in the APNR (Supplementary Fig. 4.1). To better visualise the network, only positive relatedness values were represented. The network shows that all male individuals in the dataset are related to other sampled individuals to some degree, with some individuals being more related to more individuals than others.

We attempted to measure the correlation coefficient between the social (data not shown) and genetic networks for three types of observations. In all cases, the p-values were above 0.05 (Table 4.1), and the observed coefficient never fell outside the 95% confidence interval (Fig. 4.2). As such, there is no significant correlation between sociality and genetic relatedness for the individuals analysed.

Table 4.1. Pearson's r for the correlation between genetic and social networks for the three different types of social networks. The p-values are derived from the distribution of permuted values. N: number of individuals used for the calculations.

	N	Pearson's correlation coefficient	P-value
All observations	16	0.122	0.111
Male-only groups	13	0.157	0.098
Non-musth observations	15	0.104	0.190

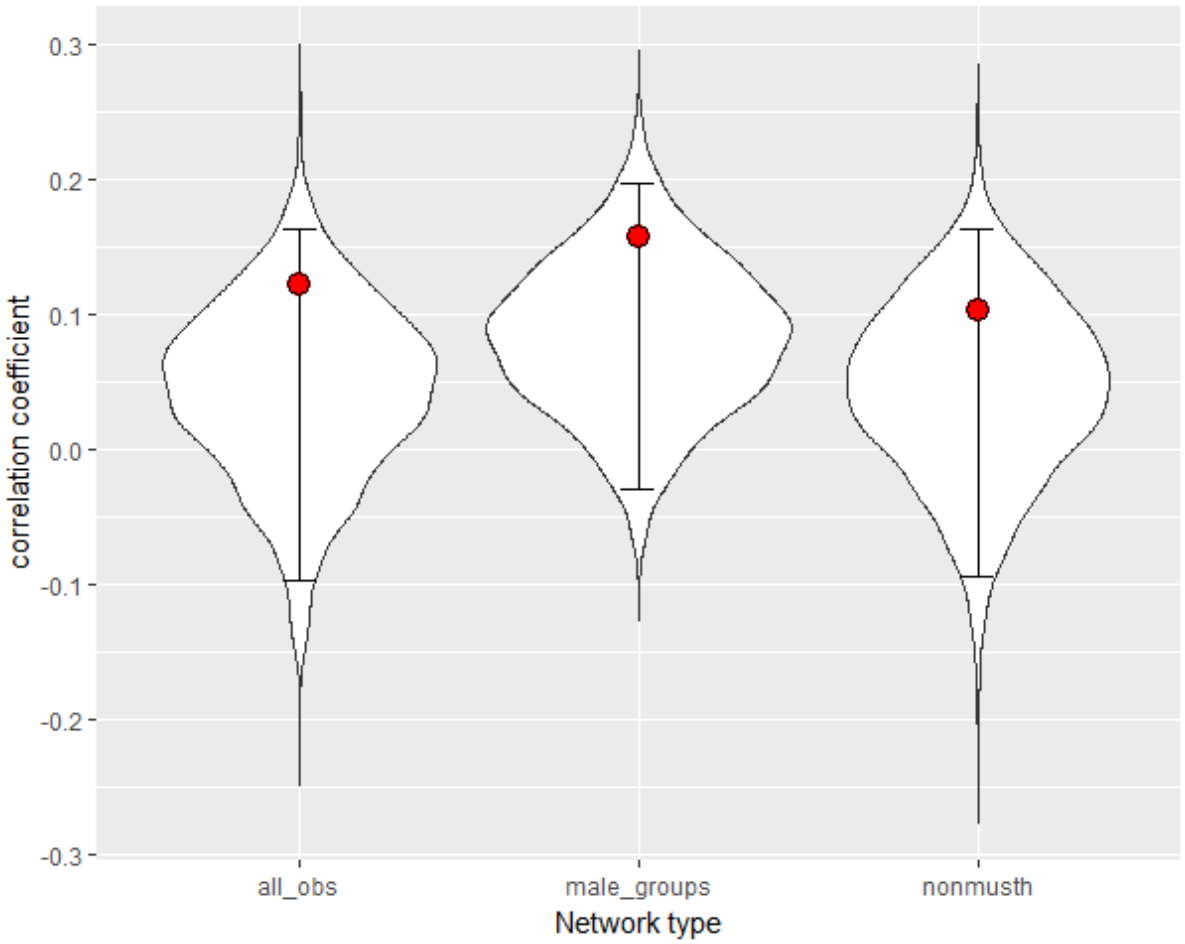


Figure 4.2. Violin plots for the correlation between social and genetic networks in relation to the distribution of the bootstrapped correlation coefficients for each network type (all observations, male-only groups, and non-musth observations). The observed correlation coefficient is signalled in red, and the 95% confidence interval is represented by the vertical bars.

4.4. Discussion

A first analysis of the impact of genetic relatedness on individuals' aggregation into groups showed that, at close distances, males are more related than expected, similarly to what is observed for females. This is a first indication that related individuals have at least partially overlapping home ranges. An analysis at a finer scale is necessary to determine whether this is a spurious result and, if not, what is the extent of such an overlap. Indeed, it would be interesting to assess if related males use the same home ranges more often than expected by chance. It is known that male elephants only avoid crossing another male's home range during musth[111], yet no study has hitherto attempted to fully study the implications of kinship on spatial use by males either during or out of musth, nor whether males are more permissive of home range overlap during musth if it is done by a related individual. To do so, it is necessary to both obtain data on home ranges and build social networks.

The study of social networks of wild animals necessitates a long study period to both observe a great number of individuals, and to observe each individual enough times to infer the workings of such a network. We did not find a correlation between a male's choice of associates and genetic relatedness. However, thus far, we have only been able to obtain enough observations to study this correlation for a small number of animals. More sightings of more males will be obtained to revisit this issue in the future.

Only two studies to date have looked at the impacts of genetic relatedness on sociality for male elephants[77,103]. They obtained conflicting results, which may be partly explained by the fact that the sampling size and methods used differ. One calculated the association index for 47 males that had been observed at least fifteen times[103], whilst the other detected 54 males and did not calculate an association index[77]. As such, for the latter study, there is no measure of how strong the observed associations were, nor is it specified how often each bull was observed. There is great variation in the duration and cohesion of male associations. If the bulls were observed few times, it is possible that their preferred associates were not sampled.

There are several ways to measure an association index. The method employed by Chiyo *et al*.[103] used a Mantel test which randomises data after building a network, whereas the method we are using randomises raw observational data before network creation. The latter method has been found to be more accurate for simulations where the 'real' result is known[119]. Therefore, once more observational data is processed, we expect to be able to provide a more accurate answer to this question, as well as study other ways in which kinship could impact sociality. For example, we aim to assess whether genetic relatedness impacts community structure as seen in female eastern grey kangaroos (*Macropus giganteus*[120]), or if related individuals occupy similar positions in a social network as observed in female elephants[121].

Associating with kin confers fitness benefits[122], but kinship is not the only factor to impact fitness, nor the only aspect influencing association. Male chimpanzees (*Pan troglodytes schweinfurthii*) increase their reproductive fitness through coalitionary aggression, particularly when associating with partners that did not form coalitions with each other[123]. In male Assamese macaques (*Macaca assamensis*) bonding with a dominant male helped to raise an individual's social rank and, consequently, improve mating chances[124]. Elephant societies are highly complex. As such, it is possible that male elephants use a panoply of information other than kinship when choosing associates. By continuing and developing this study in the future, we hope to better understand male elephant associations.

5. Final Discussion

Contrary to our original hypotheses, we found no evidence for a genetic bottleneck, a low level of genetic diversity or significant inbreeding, and we estimated a relatively high value of effective population size. We also found little genetic differentiation either within or between the two studied populations. Finally, we did not find a correlation between social bonds and genetic relatedness in males, but we obtained preliminary evidence that suggests that related individuals, regardless of sex, demonstrate at least a partial home range overlap. Both of these last results require confirmation in future studies with a higher number of observations for more males and the obtainment of home range data, respectively. Importantly, the results suggest a good genetic health of the APNR population, which makes it suitable for the obtainment of social networks representative of an undisturbed population.

A study estimated that 17,433 elephants inhabit South Africa[125]. The last census of elephants in the APNR was conducted in 2012, when it was estimated that circa 1,500 elephants occupied the area, and the highest number of individuals observed there over a one-year period was of about 2,000 in 2006. In this same year, the Kruger National Park elephant population neared 13,000 individuals[126]. The Ne/Nc ratio for the APNR is circa 0.2, above the observed value for most wild populations[60]. However, when we consider the genetic structure results, which suggest that Kruger and APNR elephants constitute a panmictic population, and pool together the censuses for 2006, this ratio declines to about 0.03. This value decreases further if we perform this calculation for the 2016 Kruger census instead. Ideally, more individuals from the Kruger and other neighbouring areas need to be sampled to allow a more reliable estimation of this parameter.

Although we did not detect a genetic bottleneck, previous studies have shown that known bottlenecks are not always detected[37], particularly when there is a high value of post-bottleneck Ne[54], as was the case here. Moreover, a small number of migrants may suffice to mask the genetic signature of bottlenecks[62,63]. Migration has not been quantified in the APNR, but a high number of migrants and translocations into the Kruger Park were reported in the 1960s[127]. Since fences between the APNR and Kruger Park were removed in the 1990s and movement occurs between them, enough migration might have occurred to conceal the signature of a bottleneck.

Migration associated with male-mediated gene flow may also explain the observed lack of genetic structure, which is similar to what has been found in other elephant populations[66,93]. However, there may still be structure at the mtDNA level due to the females' strong site fidelity[76]. A recent study found two nuclear DNA clusters across the Kruger Park (one in the north and one in the south), which showed low differentiation between them (Fst=0.022)[64]. They also found five mtDNA haplotypes across the same area, all of which were also found outside of the Kruger Park, hence suggesting that the Kruger population increased through migration and gene flow. The APNR is contiguous to the middle area of the Kruger Park, and may have been populated through a similar mechanism.

The lack of genetic differences between populations increases the risk of losing alleles to genetic drift as the probability of alleles becoming fixed increases[68]. It is necessary, therefore, to ensure that social groups and male migration are maintained. One of the ways to do so is by thoroughly controlling hunting activity. As little is known about male sociality, it is preferable to be cautious when establishing hunting quotas or when choosing which individuals can be hunted to minimise social and genetic impacts on the population. A maximum of ten bulls over 35 years of age can be hunted without causing social instability in the Greater Mapungubwe Transfrontier Conservation Area[128]. It would be of great interest to perform a similar study for the APNR and Kruger Park populations to ascertain whether current hunting quotas meet sustainable practices.

Despite having a stable population of elephants, it has been estimated that South Africa, in particular the Kruger Park area, has a deficit of 47,999 to 36,000 elephants compared to ecological benchmarks[23]. The current population needs to at least triplicate to reach this benchmark[125], despite

unfounded claims that the Kruger Park has a carrying capacity of only 7,000 elephants[126]. The combination of an increase in poaching, habitat fragmentation, unsustainable hunting quotas and human-wildlife conflict threatens to undo the rise of elephant numbers in South Africa. It is thus recommended to continue the genetic monitoring of the APNR elephants.

Future work would benefit from the amplification of mtDNA sequences to determine if there is genetic structure at the mtDNA level. Moreover, determining if female herds share the same haplotypes could help to ascertain if these groups are disturbed due to poaching[17]. This information could be helpful for studies of male sociality, as it would offer a proxy for male disturbance, which needs to be measured to ensure that the social bonds are that of an undisrupted population. It would also be of interest to assess if genetic differentiation can be found in areas of the genome that are under selection, such as genomic areas associated with tusk size.

We found evidence that individuals with higher genetic relatedness share at least partial home ranges. It would be interesting to calculate the home ranges of related individuals to better determine whether this is a valid result and, if so, the extent of this home range overlap. Such calculations should also take into account the movements of males of different ages during and out of musth, as age and musth status influence elephant behaviour and sociality[103,104,111]. We also aim to process more observational data for more bulls to better understand the impacts of genetic relatedness on male sociality and community structure, as the quantity of observations obtained thus far does not permit the attainment of robust conclusions.

Overall, this study emphasizes the importance of gene flow to maintain genetic diversity and counteract the genetic effects of culling and poaching. As such, the maintenance of connectivity between national parks and nature reserves is paramount to ensure migration between populations.

References

1. Podgórski,T., Lusseau,D., Scandura,M., Sönnichsen,L. & Jędrzejewska,B. Long-lasting, kin-directed female interactions in a spatially structured wild boar social network. *PLoS ONE* **9,** e99875 (2014).
2. Peterman,W.E. *et al.* Differential dispersal shapes population structure and patterns of genetic differentiation in two sympatric pond breeding salamanders. *Conserv. Genet.* **16,** 59–69 (2015).
3. Huchard,E., Schliehe-Diecks,S., Kappeler,P.M. & Kraus,C. The inbreeding strategy of a solitary primate, *Microcebus murinus*. *J. Evol. Biol.* **30,** 128–140 (2017).
4. Banks,S.C., Piggott,M.P., Stow,A.J. & Taylor,A.C. Sex and sociality in a disconnected world: a review of the impacts of habitat fragmentation on animal social interactions. *Can. J. Zool.* **85,** 1065–1079 (2007).
5. Winiarski,J.M., Moorman,C.E., Carpenter,J.P. & Hess,G.R. Reproductive consequences of habitat fragmentation for a declining resident bird of the longleaf pine ecosystem. *Ecosphere* **8,** e01898 (2017).
6. Cattarino,L., McAlpine,C.A. & Rhodes,J.R. Spatial scale and movement behaviour traits control the impacts of habitat fragmentation on individual fitness. *J. Anim. Ecol.* **85,** 168–177 (2016).
7. Goossens,B. *et al.* Habitat fragmentation and genetic diversity in natural populations of the Bornean elephant: Implications for conservation. *Biol. Conserv.* **196,** 80–92 (2016).
8. Gobush,K.S., Mutayoba,B.M. & Wasser,S.K. Long-term impacts of poaching on relatedness, stress physiology, and reproductive output of adult female African elephants. *Conserv. Biol.* **22,** 1590–1599 (2008).
9. Ishengoma,D.R.S. *et al.* Effects of poaching on bull mating success in a free ranging African elephant (*Loxodonta africana*) population in Tarangire National Park, Tanzania. *Conserv. Genet.* **9,** 247–255 (2008).
10. Couzin,I.D. Behavioral ecology: social organization in fission–fusion societies. *Curr. Biol.* **16,** R169–R171 (2006).
11. Blanc,J. *Loxodonta africana. The IUCN Red List of Threatened Species 2008: e.T12392A3339343*. (2008).
12. Douglas-Hamilton,I. African elephants: population trends and their causes. *Oryx* **21,** 11 (1987).
13. Buchmann,C.M., Schurr,F.M., Nathan,R. & Jeltsch,F. Habitat loss and fragmentation affecting mammal and bird communities—The role of interspecific competition and individual space use. *Ecol. Inform.* **14,** 90–98 (2013).
14. Haag,T. *et al.* The effect of habitat fragmentation on the genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (*Panthera onca*). *Mol. Ecol.* **19,** 4906–4921 (2010).
15. Archie,E.A., Morrison,T.A., Foley,C.A.H., Moss,C.J. & Alberts,S.C. Dominance rank relationships among wild female African elephants, *Loxodonta africana*. *Anim. Behav.* **71,** 117–127 (2006).
16. Ramos-Fernández,G. & Morales,J.M. Unraveling fission-fusion dynamics: how subgroup properties and dyadic interactions influence individual decisions. *Behav. Ecol. Sociobiol.* **68,** 1225–1235 (2014).
17. Archie,E.A. & Chiyo,P.I. Elephant behaviour and conservation: social relationships, the effects of poaching, and genetic tools for management. *Mol. Ecol.* **21,** 765–778 (2012).
18. Galanti,V., Preatoni,D., Martinoli,A., Wauters,L.A. & Tosi,G. Space and habitat use of the African elephant in the Tarangire–Manyara ecosystem, Tanzania. *Mamm. Biol. - Z. Für Säugetierkd.* **71,** 99–114 (2006).
19. Viljoen,P.J. Spatial distribution and movements of elephants (*Loxodonta africana*) in the northern Namib Desert region of the Kaokoveld, South West Africa/Namibia. *J. Zool.* **219,** 1–19 (1989).

20. Gobush,K.S. & Wasser,S.K. Behavioural correlates of low relatedness in African elephant core groups of a poached population. *Anim. Behav.* **78,** 1079–1086 (2009).
21. Lee,P.C., Sayialel,S., Lindsay,W.K. & Moss,C.J. African elephant age determination from teeth: validation from known individuals. *Afr. J. Ecol.* **50,** 9–20 (2011).
22. Wittemyer,G., Barner Rasmussen,H. & Douglas-Hamilton,I. Breeding phenology in relation to NDVI variability in free-ranging African elephant. *Ecography* **30,** 42–50 (2007).
23. Robson,A. S. *et al.* Savanna elephant numbers are only a quarter of their expected values. *PLOS ONE* **12,** e0175942 (2017).
24. Henley,M.D. Report on elephant movements in relation to water and the effect of the 2012 floods within the Associated Private Nature Reserves. *Unpubl. Rep. Assoc. Priv. Nat. Reserv.* **32,** (2014).
25. Ishida,Y. *et al.* Genetic connectivity across marginal habitats: the elephants of the Namib Desert. *Ecol. Evol.* **6,** 6189–6201 (2016).
26. Rees,H.C., Maddison,B.C., Middleditch,D.J., Patmore,J.R.M. & Gough,K. C. The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *J. Appl. Ecol.* **51,** 1450–1459 (2014).
27. Ishida,Y., Georgiadis,N.J., Hondo,T. & Roca,A.L. Triangulating the provenance of African elephants using mitochondrial DNA. *Evol. Appl.* **6,** 253–265 (2013).
28. Chiyo,P.I., Moss,C.J., Archie,E.A., Hollister-Smith,J.A. & Alberts,S.C. Using molecular and observational techniques to estimate the number and raiding patterns of crop-raiding elephants. *J. Appl. Ecol.* **48,** 788–796 (2011).
29. Selkoe,K.A. & Toonen,R.J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9,** 615–629 (2006).
30. Städele,V. & Vigilant,L. Strategies for determining kinship in wild populations using genetic data. *Ecol. Evol.* **6,** 6107–6120 (2016).
31. Snyder-Mackler,N. *et al.* Efficient Genome-Wide Sequencing and Low-Coverage Pedigree Analysis from Noninvasively Collected Samples. *Genetics* **203,** 699–714 (2016).
32. Vartia,S. *et al.* A novel method of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. *R. Soc. Open Sci.* **3,** 150565 (2016).
33. Bérénos,C., Ellis,P.A., Pilkington,J.G. & Pemberton,J.M. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Mol. Ecol.* **23,** 3434–3451 (2014).
34. Amorim,A. & Pereira,L. Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs. *Forensic Sci. Int.* **150,** 17–21 (2005).
35. Okello,J.B.A. *et al.* Effective population size dynamics reveal impacts of historic climatic events and recent anthropogenic pressure in African elephants. *Mol. Ecol.* **17,** 3788–3799 (2008).
36. Wittemyer,G. *et al.* Where sociality and relatedness diverge: the genetic basis for hierarchical social organization in African elephants. *Proc. R. Soc. B Biol. Sci.* **276,** 3513–3521 (2009).
37. Whitehouse,A.M. & Harley,E.H. Post-bottleneck genetic diversity of elephant populations in South Africa, revealed using microsatellite analysis. *Mol. Ecol.* **10,** 2139–2149 (2001).
38. Gobush,K., Kerr,B. & Wasser,S. Genetic relatedness and disrupted social structure in a poached population of African elephants. *Mol. Ecol.* **18,** 722–734 (2009).
39. Chiyo,P.I., Obanda,V. & Korir,D.K. Illegal tusk harvest and the decline of tusk size in the African elephant. *Ecol. Evol.* **5,** 5216–5229 (2015).
40. Steenkamp,G., Ferreira,S.M. & Bester,M.N. Tusklessness and tusk fractures in free-ranging African savanna elephants (*Loxodonta africana*). *J. S. Afr. Vet. Assoc.* **78,** 75–80 (2007).
41. Comstock,K.E., Wasser,S.K. & Ostrander,E.A. Polymorphic microsatellite DNA loci identified in the African elephant (*Loxodonta africana*). *Mol. Ecol.* **9,** 1004–1006 (2000).
42. Comstock,K.E. *et al.* Patterns of molecular genetic variation among African elephant populations. *Mol. Ecol.* **11,** 2489–2498 (2002).

43. Ahlering,M.A., Hailer,F., Roberts,M.T. & Foley,C. A simple and accurate method to sex savannah, forest and Asian elephants using noninvasive sampling techniques. *Mol. Ecol. Resour.* **11,** 831–834 (2011).
44. Kalinowski,S.T., Taper,M.L. & Marshall,T.C. Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16,** 1099–1106 (2007).
45. Rousset,F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8,** 103–106 (2008).
46. Van Oosterhout,C., Hutchinson,W.F., Wills,D.P.M. & Shipley,P. micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4,** 535–538 (2004).
47. Eggert,L.S., Ramakrishnan,U., Mundy,N.I. & Woodruff,D.S. Polymorphic microsatellite DNA markers in the African elephant (*Loxondonta africana*) and their use in the Asian elephant (*Elephas maximus*). *Mol. Ecol.* **9,** 2223–2225 (2000).
48. Archie,E.A., Moss,C.J. & Alberts,S.C. Characterization of tetranucleotide microsatellite loci in the African savannah elephant (*Loxodonta africana africana*). *Mol. Ecol. Notes* **3,** 244–246 (2003).
49. Valière,N. GIMLET: a computer program for analysing genetic individual identification data. *Mol. Ecol. Resour.* **2,** 377–379 (2002).
50. Kalinowski,S.T., Wagner,A.P. & Taper,M.L. ml-relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol. Ecol. Notes* **6,** 576–579 (2006).
51. Do,C. *et al.* NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Resour.* **14,** 209–214 (2014).
52. Piry,S., Luikart,G. & Cornuet,J.M. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *J. Hered.* **90,** 502–503 (1999).
53. Luikart,G., Allendorf,F.W., Cornuet,J.M. & Sherwin,W.B. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* **89,** 238–247 (1998).
54. Peery,M.Z. *et al.* Reliability of genetic bottleneck tests for detecting recent population declines. *Mol. Ecol.* **21,** 3403–3418 (2012).
55. England,P.R., Cornuet,J.M., Berthier,P., Tallmon,D.A. & Luikart,G. Estimating effective population size from linkage disequilibrium: severe bias in small samples. *Conserv. Genet.* **7,** 303–308 (2006).
56. Belkhir,K., Borsa,P., Chikhi,L. & Bonhomme,F. *Genetix 4.05: Windows^{TM} Software for Population Genetics*. (University of Montpelier II, Laboratoire Genome de Populations, Montpelier, France, 1996).
57. Kalinowski,S.T. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* **5,** 187–189 (2005).
58. Narum,S.R. Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conserv. Genet.* **7,** 783–787 (2006).
59. Waples,R.S. & Do,C. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol. Appl.* **3,** 244–262 (2010).
60. Frankham,R. Effective population size/adult population size ratios in wildlife: a review. *Genet. Res.* **66,** 95–107 (1995).
61. Archie,E.A. *et al.* Fine-scale population genetic structure in a fission-fusion society. *Mol. Ecol.* **17,** 2666–2679 (2008).
62. Cornuet,J.M. & Luikart,G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144,** 2001–2014 (1996).
63. Keller,L.F. *et al.* Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. *Proc. R. Soc. B Biol. Sci.* **268,** 1387–1394 (2001).

64. de Flamingh,A., Roca,A.L. & van Aarde,R.J. Origin and phylogeography of African savannah elephants (*Loxodonta africana*) in Kruger and nearby parks in southern Africa. *Conserv. Genet.* (2017).
65. Ahlering,M.A. *et al.* Identifying source populations and genetic structure for savannah elephants in human-dominated landscapes and protected areas in the Kenya-Tanzania borderlands. *PLoS ONE* **7,** e52288 (2012).
66. Okello,J.B.A. *et al.* Population genetic structure of savannah elephants in Kenya. *J. Hered.* **99,** 443–452 (2008).
67. Wasser,S.K. *et al.* Assigning African elephant DNA to geographic region of origin: applications to the ivory trade. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 14847–14852 (2004).
68. Dobson,F.S., Chesser,R.K., Hoogland,J.L., Sugg,D.W. & Foltz,D.W. The influence of social breeding groups on effective population size in black-tailed prairie dogs. *J. Mammal.* **85,** 58–66 (2004).
69. Evanno,G., Regnaut,S. & Goudet,J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14,** 2611–2620 (2005).
70. Paetkau,D., Slade,R., Burden,M. & Estoup,A. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* **13,** 55–65 (2004).
71. Western,D. & Lindsay,W.K. Seasonal herd dynamics of a savanna elephant population. *Afr. J. Ecol.* **22,** 229–244 (1984).
72. Graham,M.D., Douglas-Hamilton,I., Adams,W.M. & Lee,P.C. The movement of African elephants in a human-dominated land-use mosaic. *Anim. Conserv.* **12,** 445–455 (2009).
73. Shannon,G., Matthews,W.S., Page,B.R., Parker,G.E. & Smith,R.J. The effects of artificial water availability on large herbivore ranging patterns in savanna habitats. *Divers. Distrib.* **15,** 776–783 (2009).
74. Eggert,L.S., Rasner,C.A. & Woodruff,D.S. The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proc. R. Soc. Lond. B Biol. Sci.* **269,** 1993–2006 (2002).
75. Lorenzen,E.D., Heller,R. & Siegismund,H.R. Comparative phylogeography of African savannah ungulates. *Mol. Ecol.* **21,** 3656–3670 (2012).
76. Nyakaana,S., Arctander,P. & Siegismund,H.R. Population structure of the African savannah elephant inferred from mitochondrial control region sequences and nuclear microsatellite loci. *Heredity* **89,** 90–98 (2002).
77. Ahlering,M.A., Maldonado,J.E., Fleischer,R.C., Western,D. & Eggert,L.S. Fine-scale group structure and demography of African savanna elephants recolonizing lands outside protected areas. *Divers. Distrib.* **18,** 952–961 (2012).
78. Guillot,G., Mortier,F. & Estoup,A. GENELAND: a computer package for landscape genetics. *Mol. Ecol. Resour.* **5,** 712–715 (2005).
79. Pritchard,J.K., Stephens,M. & Donnelly,P. Inference of population structure using multilocus genotype data. *Genetics* **155,** 945–959 (2000).
80. Chen,C., Durand,E., Forbes,F. & François,O. Bayesian clustering algorithms ascertaining spatial population structure. *Mol. Ecol. Notes* **7,** 747–756 (2007).
81. Basto,M.P. *et al.* Assessing genetic structure in common but ecologically distinct carnivores: the stone marten and red fox. *PloS One* **11,** e0145165 (2016).
82. Blanchet,E. *et al.* A comparative analysis of fine-scale genetic structure in three closely related syntopic species of the grasshopper genus *Calliptamus*. *Can. J. Zool.* **90,** 31–41 (2012).
83. The Geneland development group. Population genetic and morphometric data analysis using R and the Geneland program. (2017).

84. Earl,D.A. & vonHoldt,B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4,** 359–361 (2012).
85. Jakobsson,M. & Rosenberg,N.A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23,** 1801–1806 (2007).
86. Rosenberg,N.A. distruct: a program for the graphical display of population structure. *Mol. Ecol. Notes* **4,** 137–138 (2003).
87. Guillot,G. On the inference of spatial structure from population genetics data. *Bioinformatics* **25,** 1796–1801 (2009).
88. Jachmann,H., Berry,P.S.M. & Imae,H. Tusklessness in African elephants: a future trend. *Afr. J. Ecol.* **33,** 230–235 (1995).
89. Whitehouse,A.M. Tusklessness in the elephant population of the Addo Elephant National Park, South Africa. *J. Zool.* **257,** 249–254 (2002).
90. Raubenheimer,E.J. Development of the tush and tusk and tusklessness in African elephant (*Loxodonta africana*). *Koedoe* **43,** 57–64 (2000).
91. de Flamingh,A., Sole,C.L. & van Aarde,R.J. Genetic evidence for spatial structuring in a continuous African elephant (*Loxodonta africana*) population. *Conserv. Genet.* **16,** 613–623 (2015).
92. Archie,E.A. *et al.* Behavioural inbreeding avoidance in wild African elephants. *Mol. Ecol.* **16,** 4138–4148 (2007).
93. Nyakaana,S. & Arctander,P. Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: evidence for male-biased gene flow. *Mol. Ecol.* **8,** 1105–1115 (1999).
94. FitzSimmons,N.N., Moritz,C., Limpus,C.J., Pope,L. & Prince,R. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147,** 1843–1854 (1997).
95. Clegg,S.M., Hale,P. & Moritz,C. Molecular population genetics of the red kangaroo (*Macropus rufus*): mtDNA variation. *Mol. Ecol.* **7,** 679–686 (1998).
96. Allendorf,F.W., Hohenlohe,P.A. & Luikart,G. Genomics and the future of conservation genetics. *Nat. Rev. Genet.* **11,** 697–709 (2010).
97. Pampoulie,C. *et al.* The genetic structure of Atlantic cod (*Gadus morhua)* around Iceland: insight from microsatellites, the *Pan* I locus, and tagging experiments. *Can. J. Fish. Aquat. Sci.* **63,** 2660–2674 (2006).
98. Nussey,D.H. *et al.* Rapidly declining fine-scale spatial genetic structure in female red deer. *Mol. Ecol.* **14,** 3395–3405 (2005).
99. Byrne,R. *et al.* Do elephants show empathy? *J. Conscious. Stud.* **15,** 204–225 (2008).
100. McComb,K., Moss,C., Sayialel,S. & Baker,L. Unusually extensive networks of vocal recognition in African elephants. *Anim. Behav.* **59,** 1103–1109 (2000).
101. Douglas-Hamilton,I., Bhalla,S., Wittemyer,G. & Vollrath,F. Behavioural reactions of elephants towards a dying and deceased matriarch. *Appl. Anim. Behav. Sci.* **100,** 87–102 (2006).
102. Vance,E.A., Archie,E.A. & Moss,C.J. Social networks in African elephants. *Comput. Math. Organ. Theory* **15,** 273–293 (2009).
103. Chiyo,P. I. *et al.* Association patterns of African elephants in all-male groups: the role of age and genetic relatedness. *Anim. Behav.* **81,** 1093–1099 (2011).
104. Goldenberg,S.Z., de Silva,S., Rasmussen,H.B., Douglas-Hamilton,I. & Wittemyer,G. Controlling for behavioural state reveals social dynamics among male African elephants, *Loxodonta africana*. *Anim. Behav.* **95,** 111–119 (2014).
105. Evans,K.E. & Harris,S. Adolescence in male African elephants, *Loxodonta africana*, and the importance of sociality. *Anim. Behav.* **76,** 779–787 (2008).

106. Archie,E.A., Moss,C.J. & Alberts,S.C. The ties that bind: genetic relatedness predicts the fission and fusion of social groups in wild African elephants. *Proc. R. Soc. B Biol. Sci.* **273,** 513–522 (2006).
107. Petrie,M., Krupa,A. & Burke,T. Peacocks lek with relatives even in the absence of social and environmental cues. *Nature* **401,** 155–157 (1999).
108. Bradley,B.J., Doran-Sheehy,D.M., Lukas,D., Boesch,C. & Vigilant,L. Dispersed male networks in western gorillas. *Curr. Biol.* **14,** 510–513 (2004).
109. McKnight,B.L. Relationship between group dynamics and spatial distribution of African elephants in a semi-arid environment. *Afr. J. Ecol.* **53,** 439–446 (2015).
110. Grainger,M., Aarde,R. & Whyte,I. Landscape heterogeneity and the use of space by elephants in the Kruger National Park, South Africa. *Afr. J. Ecol.* **43,** 369–375 (2005).
111. Stokke,S. & du Toit,J.T. Sexual segregation in habitat use by elephants in Chobe National Park, Botswana. *Afr. J. Ecol.* **40,** 360–371 (2002).
112. Pew,J., Muir,P.H., Wang,J. & Frasier,T.R. related: an R package for analysing pairwise relatedness from codominant molecular markers. *Mol. Ecol. Resour.* **15,** 557–561 (2015).
113. Hardy,O.J. & Vekemans,X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Resour.* **2,** 618–620 (2002).
114. Queller,D.C. & Goodnight,K.F. Estimating relatedness using genetic markers. *Evolution* **43,** 258–275 (1989).
115. O'Connell-Rodwell,C.E. Keeping an 'Ear' to the Ground: Seismic Communication in Elephants. *Physiology* **22,** 287–294 (2007).
116. Wang,J. coancestry: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol. Ecol. Resour.* **11,** 141–145 (2011).
117. Bastian,M., Heymann,S., Jacomy,M. *et al*. Gephi: an open source software for exploring and manipulating networks. *Icwsm* **8,** 361–362 (2009).
118. Whitehead,H. *Analyzing animal societies: quantitative methods for vertebrate social analysis*. (University of Chicago Press, 2008).
119. Farine,D.R. A guide to null models for animal social network analysis. *Methods Ecol. Evol.* (2017).
120. Best,E.C., Seddon,J.M., Dwyer,R.G. & Goldizen,A.W. Social preference influences female community structure in a population of wild eastern grey kangaroos. *Anim. Behav.* **86,** 1031–1040 (2013).
121. Goldenberg,S.Z., Douglas-Hamilton,I. & Wittemyer,G. Vertical transmission of social roles drives resilience to poaching in elephant networks. *Curr. Biol.* **26,** 75–79 (2016).
122. Hamilton,W.D. The Genetic Evolution of Social Behaviour. *J. Theor. Biol.* **7,** 17–52 (1964).
123. Gilby,I.C. *et al.* Fitness benefits of coalitionary aggression in male chimpanzees. *Behav. Ecol. Sociobiol.* **67,** 373–381 (2013).
124. Schülke,O., Bhagavatula,J., Vigilant,L. & Ostner,J. Social Bonds enhance reproductive success in male macaques. *Curr. Biol.* **20,** 2207–2210 (2010).
125. Chase,M.J. *et al.* Continent-wide survey reveals massive decline in African savannah elephants. *PeerJ* **4,** e2354 (2016).
126. Owen-Smith,N., Kerley,G.I.H., Page,B., Slotow,R. & Van Aarde,R.J. A scientific perspective on the management of elephants in the Kruger National Park and elsewhere. *South Afr. J. Sci.* **102,** 389–394 (2006).
127. Hall-Martin,A.J. Distribution and status of the African elephant *Loxodonta africana* in South Africa, 1652-1992. *Koedoe* **35,** 65–88 (1992).
128. Selier,S.A.J., Page,B.R., Vanak,A.T. & Slotow,R. Sustainability of elephant hunting across international borders in southern Africa: A case study of the greater Mapungubwe Transfrontier Conservation Area. *J. Wildl. Manag.* **78,** 122–132 (2014).

129. Wang,J. An estimator for pairwise relatedness using molecular markers. *Genetics* **160,** 1203–1215 (2002).
130. Li,C.C., Weeks,D.E. & Chakravarti,A. Similarity of DNA fingerprints due to chance and relatedness. *Hum. Hered.* **43,** 45–52 (1993).
131. Lynch,M. & Ritland,K. Estimation of pairwise relatedness with molecular markers. *Genetics* **152,** 1753–1766 (1999).

Supplementary Data

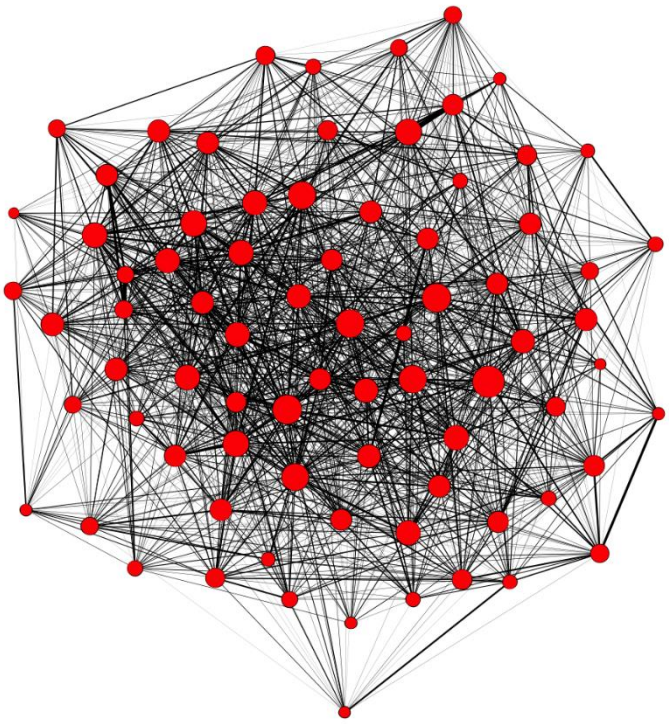
Supplementary Table 2.1. He and Ho were calculated in CERVUS and HWE was evaluated in Genepop using 20 unrelated individuals. Null allele frequencies were calculated using the Oosterhout's estimator in Micro-Checker for all samples. The dropout and false allele rates were calculated in GIMLET using a subset of 53 randomly chosen samples. N: number of alleles. He: expected heterozygosity. Ho: observed heterozygosity. HWE (A/B): Hardy-Weinberg equilibrium assessment for sets A and B of unrelated individuals. *: Markers out of HWE after Benjamini and Yekutieli's correction for multiple testing. Freq.: Frequency.

Locus	N	Allele Range	He	Ho	HWE (A/B)	Null Allele Freq.	Dropout Freq.	False Allele Freq.
FH1	5	99 – 99	0.726	0.810	0.5263 / 0.8265	0.0106	0.028	0.058
FH19	8	192 – 206	0.750	0.619	0.6102 / 0.4714	0.031	0.014	0
FH39	12	238 – 278	0.786	0.714	0.7357 / 0.549	-0.0297	0	0.028
FH40	6	246 – 256	0.473	0.381	0.4975 / 0.1303	-0.0585	0.022	0.019
FH48	8	178 – 192	0.660	0.714	0.2032 / 0.5374	0.0051	0.043	0
FH60	4	154 – 160	0.418	0.524	0.5564 / 0.7274	0.013	0	0
FH67	8	101 – 115	0.685	0.429	0.8705 / 0.0025*	0.0365	0	0.038
FH71	3	71 – 75	0.589	0.571	0.4944 / 0.8852	0.0541	0.013	0
FH94	6	226 – 238	0.747	0.714	0.2956 / 0.3936	0.0031	0.023	0.028

FH103	5	154 – 162	0.538	0.333	0.2361 / 0.0213	-0.0005	0.016	0
LA5	7	147 – 161	0.626	0.571	0.1662 / 0.1235	-0.0173	0.045	0
LA6	5	166 – 182	0.408	0.524	0.772 / 0.4612	-0.0355	0.029	0.049
LaT06	13	283 – 401	0.715	0.429	0.3142 / 0.0084*	0.1204	0.099	0.103
LaT08	14	183 – 235	0.844	0.905	0.9522 / 0.1446	-0.0134	0.005	0.095
LaT13	11	229 – 269	0.822	0.667	0.0069* / 0.0462	0.0405	0.033	0
LaT18	10	290 – 324	0.827	0.810	0.5845 / 0.7996	0.0454	0.063	0.216
LaT24	13	202 – 258	0.846	0.810	0.6947 / 0.9418	0.028	0.027	0.033
LaT25	8	300 – 328	0.739	0.524	0.2053 / 0.0428	0.0849	0.102	0.189

Supplementary Table 4.1. Correlation coefficients between observed and expected values of relatedness for four different relationships (full siblings, half siblings, parent-offspring, unrelated) using four different relatedness estimators.

Estimator	100 simulated relationships	500 simulated relationships
Wang[129]	0.832	0.836
Li *et al.*[130]	0.834	0.829
Lynch & Ritland[131]	0.784	0.781
Queller & Goodnight[114]	0.842	0.836



Supplementary Figure 4.1. Genetic network of adult males (over 20 years old) from APNR (n = 78). Each node represents an individual and the edges represent the pairwise genetic relatedness between connecting nodes. The larger the nodes, the higher the degree of the node (i.e. it has a higher number of connections to other nodes). The thicker the edges, the higher the relatedness between individuals.